

Public Perceptions about COVID-19 Vaccines: Policy Implications from US Spatiotemporal Sentiment Analytics

G. G. MD. NAWAZ ALI¹ (MEMBER, IEEE), MD. MOKHLESUR RAHMAN^{2, 8}, MD. AMJAD HOSSAIN³ (MEMBER, IEEE), MD. SHAHINOOR RAHMAN⁴, KAMAL CHANDRA PAUL⁵ (GRADUATE STUDENT MEMBER, IEEE), JEAN-CLAUDE THILL⁶, AND JIM SAMUEL⁷

¹Department of Applied Computer Science, University of Charleston, WV 25304, USA

²The William States Lee College of Engineering, University of North Carolina at Charlotte, NC 28223, USA

³Department of Computer Science, Math and Engineering, Shepherd University, WV 25443, USA

⁴Department of Earth and Environmental Sciences, New Jersey City University, NJ 07305, USA

⁵Department of Electrical and Computer Engineering, University of North Carolina at Charlotte, NC 28223, USA

⁶Department of Geography and Earth Sciences, School of Data Science, University of North Carolina at Charlotte, NC 28223, USA

⁷College of Business, University of Charleston, WV 25304, USA

⁸Department of Urban and Regional Planning, Khulna University of Engineering & Technology (KUET), Khulna-9203, Bangladesh

Corresponding author: G. G. Md. Nawaz Ali (e-mail: ggmdnawazali@ucwv.edu).

ABSTRACT There exists a compelling need to better understand the temporal dynamics of public sentiment towards COVID-19 vaccines in the US on a national and state-wise level for facilitating appropriate public policy applications. Our analysis of social media data from early February of 2021 and late March of 2021 shows that in spite of overall strength of positive sentiment, and increasing numbers of Americans being fully vaccinated, negative sentiment about COVID-19 vaccines still persists among sections of people who are hesitant towards the vaccine. In this study, we performed sentiment analytics on vaccine tweets, studied changes in public sentiment over time, conducted vaccination sentiment validation using actual vaccination data from the US CDC and Household Pulse Survey (HPS), explored influence of maturity of Twitter user-accounts and generated geographic mapping of sentiments by location of Twitter users. Furthermore, we leverage the emotion polarity based Public Sentiment Scenarios (PSS) framework which was developed for COVID-19 sentiment analytics, to systematically analyze directions for public policy processes to potentially improve the administration of vaccines. Application of the PSS framework provides important time sensitive insights for state and federal government agencies and associated organizations to better implement public policy processes for healthcare management, communication, transparency, motivation and societal operational policies such as social distancing. These insights are expected to contribute to processes that can expedite the vaccination program and move closer to the cherished herd immunity goal.

INDEX TERMS Vaccine, Sentiment analysis, Public Sentiment Scenarios framework, COVID-19, Coronavirus, Twitter, Textual analytics, Public policy

I. INTRODUCTION

The novel coronavirus disease 2019 (COVID-19) has already claimed the lives of more than 582,263 Americans, and the numbers continue to rise, albeit at a variable pace (May 15, 2021 - [1]). The COVID-19 pandemic is a crucial health crisis in the history of humankind with more than 162.9 million confirmed cases and 3.3 million deaths as of May 15, 2021 [2], which has had tremendous disruptive effects on public and mental health, social structures, and local and global economies in almost every country of the world [3]–

[5]. Vaccines have the potential to catalyze the development of herd immunity¹ and stop the viral COVID-19 rampage. According to early stage research, to achieve herd immunity in US, 58%-85% adults need to be fully vaccinated with a 95% vaccination efficacy². Given the estimate that vac-

¹Herd immunity is a form of indirect protection from an infectious disease when a sufficient number of people of a community became immune to an infection [6].

²Two mRNA vaccines: Pfizer-BioNTech and Moderna have at most 95% efficacy in clinical trials [7], [8].

inations can only reduce 75% of transmissions, achieving herd immunity will require higher proportions (78%-94%) of the population to be vaccinated [9]. So far (as of May 15, 2021), according to the US Centers for Disease Control and Prevention (US CDC), 46.6% of US adults (age ≥ 18) are fully vaccinated. However, interestingly the number of daily administered doses are falling (Fig. 1) and millions of Americans appear to be forfeiting or indefinitely postponing the second round of vaccine doses [10]. A recent survey estimated that 40% adults are hesitant to take COVID-19 vaccine [11]. This huge number of vaccine-hesitant adults might be detrimental to the vaccine induced herd immunity objective.

In this research study, we identify public sentiment towards vaccines using Twitter data. Our research studies the progression of public sentiment towards vaccine from early February of 2021 to late March of 2021 in the whole US as well as state level. Insights from Twitter data were verified and validated by the actual vaccination data collected from CDC [12] and Nationwide Household Pulse Survey (HPS) conducted by the US Census Bureau from April 14 to April 26, 2021 [13]. The in-principle value and findings of this research have significant implications for state and federal government public policy formation processes to expedite the vaccination program and achieve the cherished herd immunity goal.

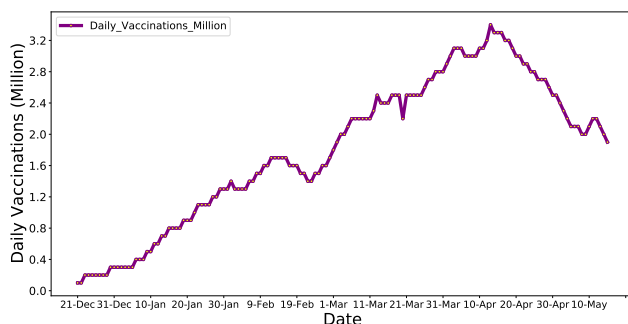


FIGURE 1: Total number of Daily vaccination (million) in the US (Data source: [12]).

Public policy has played a dominant role in addressing the COVID-19 crisis. The success experienced by countries across the world has depended on the effectiveness of their COVID-19 public policies pertaining to healthcare, communication, education, motivation and societal operational policies such as social distancing. Due to the nonexistence of the COVID-19 vaccine in the early stages of the pandemic, public policies focused on different non-pharmaceutical interventions (e.g., lockdown, restrictions on mass gathering, travel ban, border closing, testing and contact tracing) and economic stimuli (e.g., donation, loan, debt relief) have been implemented to contain COVID-19 pandemic and mitigate associated risks [14]–[18]. Simultaneously, scientists and researchers have worked on developing COVID-19 vaccines, which are critical to control virus diffusion, protecting hu-

man lives, and social and economic recovery [14], [16]. The success of COVID-19 vaccine production, distribution, and actual administration is significantly dependent on the public acceptance of the vaccine amid viral misinformation, hesitancy, fear of side effects, and allergies [19]–[22]. Considering the urgent need, this study has intended to evaluate the sentiment of the people on the COVID-19 vaccine and propose some public policies for state and federal government for expediting the COVID-19 vaccination.

To reach the herd immunity through the vaccinations, there are two phases: (1) ensuring the availability of vaccines for all, and (2) administering the vaccine. Having enough vaccine available for the population, the biggest challenge is to administer the vaccine to enough population because of many common issues including vaccine hesitancy. Usually, developing a vaccine is time and resource-intensive. Hence, immediate availability of vaccines creates consumers concern about safety and effectiveness [14]. Moreover, vaccine hesitancy is influenced by the level of confidence (i.e., lack of trust in vaccine, provider, and regulatory authority), complacency (i.e., no perceived need for a vaccine and do not value the vaccine), and convenience (i.e., difficult access to vaccines, cost) [14], [23]–[26]. Researchers have found that COVID-19 vaccine hesitancy is very high in many countries of the world due to misinformation about side effects, hasty production, lack of trust, and doubt in efficiency [23], [27]. Thus, identifying the reasons and information for vaccine hesitancy is an important research direction for increasing the COVID-19 vaccine uptake.

The public sentiment analysis based on social media data has become the de facto technique to identify and address many trending issues of our society. [4], [5], [28]–[30]. In this research work, we have analyzed the public sentiment about COVID-19 vaccine using Twitter data (tweets that are posted from different locations of the US). We have analyzed the progression of public sentiment about vaccination by analyzing the tweets that are posted on the early week of Feb'21 and last week of March'21. After cleaning and preprocessing, we have taken 5,131 and 9,036 tweets for Feb'21 and March'21, respectively, for sentiment analysis. The characteristics of those tweets are displayed in Fig. 2. Using `sentimentR` package, we have classified the tweets into three sentiment groups: positive sentiment, negative sentiment and neutral sentiment. Additionally, to validate the sentiment score computation, we have also computed the sentiment score of the same set of tweets using two lexicon-based Python libraries and found the similar results. Overall, we have found the dominant positive sentiment over negative sentiment about vaccination in the US in the both time frames as shown in Fig. 3. However, interestingly, we have seen that as time progresses, positive sentiment declines (by around 9.5%) and neutral sentiment increases (by almost 9%). That means just giving more time to take the vaccination is not going to work in the US to reach the herd immunity goal. Another important finding from this study reveals that sentiment varies from US one state to another. State-wise sentiment score has a

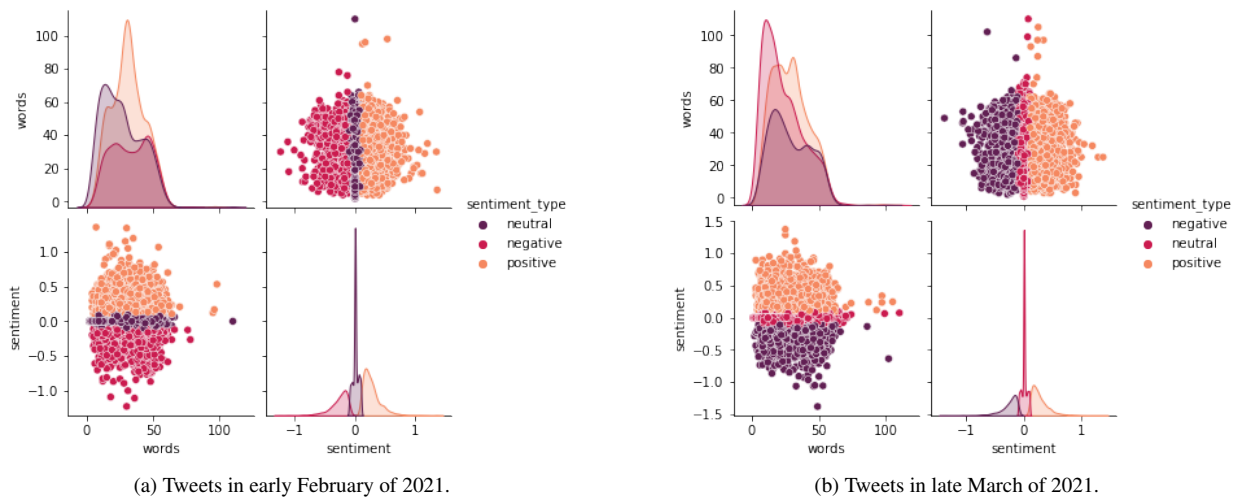


FIGURE 2: Characteristics of vaccine tweets.

proportional relation to the people willingness towards vaccination. Some states have dominant negative sentiment, and those states have lower vaccination rate as well. Hence, by devising a Public Sentiment Scenarios (PSS) framework, we recommend a number of public policies for the state and federal government and suggest to be proactive or incentivize and take state-wise different actions to encourage adult Americans to get vaccinated so that the country as a whole can reach the herd immunity goal sooner.

The rest of the paper is organized as follows. Section II discusses the related work towards vaccination sentiment, Twitter analytics, COVID-19 vaccine hesitancy, and acceptance. Section III exhibits some of the representative tweets of positive, negative, and neutral sentiments. Section IV illustrates the Public Sentiment Scenarios (PSS) framework which is used to structure hypotheses for statistical analysis, and to evaluate the public policy implications of the sentiment statistical analysis. Section V discusses about the methods of data collection from Twitter, preprocessing, computing sentiments from tweets, and spatiotemporal analysis of the sentiments. The detailed statistical, state-wise spatiotemporal analysis of the sentiments, verification and validation with the actual vaccination information, and limitation of the studies are analyzed and discussed in Section VI. Section VII discusses policy implications based on the study findings. Finally, Section VIII concludes this work with stating some future research directions.

II. RELATED WORK

A. TWITTER ANALYTICS

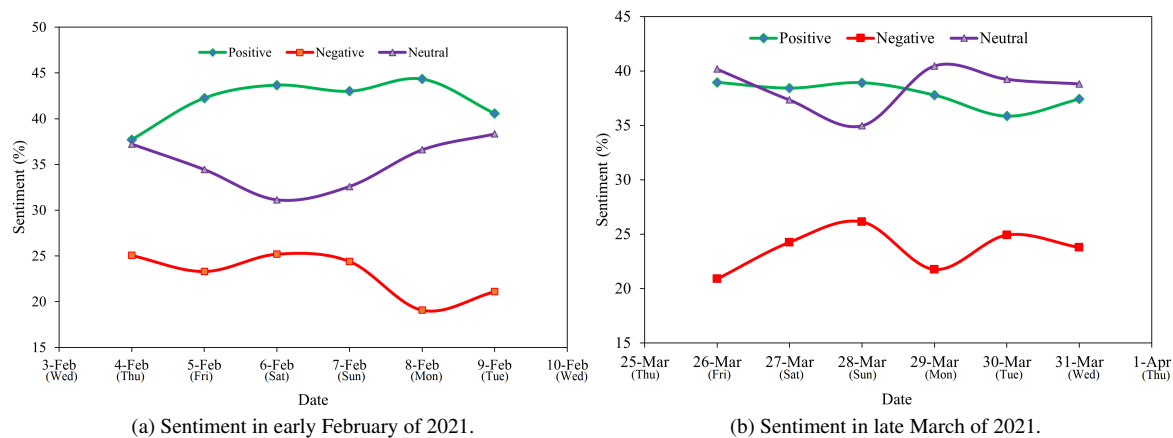
Sentiment analysis from tweets is increasingly becoming an useful and powerful tool to understand the individuals' perception about an ongoing important thing for the mankind and for the society, for example pandemic, disaster, vaccination etc. Thus, many previous have used Twitter data to understand public perceptions and associated factors that influence people's opinions on the matter of interest. For

example, Samuel *et al.* [4] investigated the insights of fear-sentiment over time from Twitter data while COVID-19 hits the US in February 2020. Similarly, just after the first wave of COVID-19 in the US, Samuel *et al.* [5] and Rahman *et al.* [18] used Twitter data to gauge public sentiment towards reopening US economy and the factors that control reopening sentiments of the people.

Some studies have also used Twitter data to apprehend public perceptions on the COVID-19 vaccine. For example, Thelwall *et al.* [28] used tweets to understand what types of vaccine hesitancy information is shared on Twitter so that appropriate actions can be taken to stop spreading the misinformation/disinformation. Lyu *et al.* [29] used Twitter data to capture public opinions on the potential COVID-19 vaccines by classifying the tweets into pro-vaccine, vaccine-hesitant, and anti-vaccine groups. Similarly, Hilary *et al.* [30] analyzed over a million vaccine-related tweets posted from the period 2011 to 2019 to get the public opinion about vaccines. They classified tweets into positive, negative and neutral towards vaccination based on geographical area and historical period. Clarivate's social intelligence experts [31] analyzed tweets to understand the evolving perspectives of Americans on COVID-19 vaccines. They observed a significant positive shift from September'20 to January'21 towards getting vaccinated. Raghupathi *et al.* [32] used natural language toolkit (NLTK) on Twitter data to explore patterns and public opinions towards measles vaccination. Salathé and Khandelwal [33] measured the evolution and distribution of sentiments towards the novel influenza A(H1N1) vaccine. Thus, Twitter data is extensively used to have knowledge of the public perceptions on any ongoing issues around the world.

B. COVID-19 VACCINE ON SOCIAL MEDIA

A few studies have investigated public perceptions and concerns of the COVID-19 vaccines in the US and a few other countries by using social media data, such as Twitter and



(a) Sentiment in early February of 2021.

(b) Sentiment in late March of 2021.

FIGURE 3: Temporal sentiment about vaccines from the tweets extracted on first week of Feb'21 and last week of March'21.

Facebook.

Lyu *et al.* [29] explored public opinion on the COVID-19 vaccine in the US after collecting data from 20,000 Twitter users. Classifying them into pro-vaccine, vaccine-hesitant, and anti-vaccine they found that 57.65% are pro-vaccine, 19.30% are vaccine-hesitant, and 23.05% are anti-vaccine. A lower number of pro-vaccine users are observed in the southeast regions of the US and clusters of Ohio, Indiana, and Kentucky. This study also found that the Americans are most concerned about the safety, effectiveness, and political issues regarding COVID-19 vaccines.

Hussain *et al.* [34] conducted a study to analyze public sentiments towards COVID-19 vaccines in US and UK by extracting information from Facebook and Twitter. More than 300,000 social media posts (i.e., 144,864 Facebook posts and 98,385 tweets from US and 23,571 Facebook posts and 40,268 tweets from UK) on COVID-19 vaccines were collected from March 1 to November 22, 2020. Using natural language processing and deep learning-based methods they analyzed public sentiments and concerns on the COVID-19 vaccines. Results indicate that about 56%, 24%, and 18% of the posts expressed positive, negative, and neutral sentiments, respectively in US. Similarly, 58%, 22%, and 17% of the social media users expressed positive, negative, and neutral sentiments, respectively in UK. In the US, positive sentiment on Facebook was prominent with a peak in Aug'20 due to progress in COVID-19 research and a negative trend achieved a peak in mid-Sept'20 due to the hasty development of the vaccine.

Janessa *et al.* [35] explored the reasons for COVID-19 vaccine hesitancy in Canada. Using the Theoretical Domains Framework they analyzed the contents of 3915 tweets based on knowledge, beliefs on consequences, environmental context and resources, social influence, and emotion. A total of 605 (15.45%) tweets are identified that showed vaccine hesitancy in different non-exclusive themes. About 48.3% of these tweets expressed concern about safety due to the rapid development of a vaccine, 32.4% were skeptical about the po-

litical motivation towards COVID-19 vaccine development, 26.3% of tweets indicated the lack of knowledge of the users due to misinformation about COVID-19, vaccination, and herd immunity. In addition, 8.4% of tweets expressed their mistrust of the vaccine due to lack of evidence of vaccine taken-up by the high government officials, company CEO etc.

Eibensteiner *et al.* [36] assessed the perceptions of people on safety and acceptance of COVID-19 vaccine worldwide using Twitter data. Creating two Twitter polls and pinning them to the Digital Health and Patient Safety Platform timeline for 1 week in mid-Feb'21, they collected data from Twitter users and influential Twitter users. The results of first poll show that about 45.9% of tweeter users agreed on the adequacy safety of the COVID-19 vaccine (i.e., all vaccines are safe). In contrast, about 41.7% were uncertain about the safety. Only 5.2% mentioned that the vaccines are generally unsafe for them. In second poll, about 82.8% users responded positively to accept vaccine and only 8% rejected to undergo vaccination.

C. COVID-19 VACCINATION SENTIMENT AND HESITANCY

The main hindrance to reach herd immunity through vaccinations is the vaccine hesitancy. So, it is necessary to study the major factors that contribute to the vaccine hesitancy in people. Conducting an online-based survey, researchers in [37] collected data from 804 adult individuals in the US and found that about 14.8% of respondents disagreed to receive COVID-19 vaccine and 23% were uncertain. The study also observed that the men, white and non-Hispanic race, college graduates, Democrats, married or partnered, people with pre-existing conditions, and persons who have up taken influenza vaccination in 2019-2020 are more likely to get vaccinated for COVID-19. CNN/MSNBC viewers have more intentions (76.4%) than Fox News viewers (57.3%) in terms of getting vaccinated.

Fisher *et al.* [38] conducted a study to assess the intention

of people to receive COVID-19 vaccine and identify the factors for vaccine hesitancy in the US. From the cross-sectional survey conducted from 16 to 20 April 2020, they collected data from 1000 adult Americans. Upon analyzing the data, they found that about 31.6% are not sure and 10.8% are not interested to undertake vaccine. The study suggested the targeted and multi-faceted efforts to increase the acceptance of the COVID-19 vaccine.

Jeremy *et al.* [39] conducted four online surveys to collect information from 5018 French people (18+) on the attitude of COVID-19 vaccines in April 2020. The study result indicated that almost 25% people refused to take the vaccine. Out of those who refused the vaccination, about 64.4% of them mentioned that the vaccine invented in rush could be too unsafe for them, 8% had no trust in vaccine, while 27.6% people did not mention any specific reasons of vaccine rejection. The study also mention that the people who are close to governing political parties have a more positive attitude toward vaccine than those who are not.

By analyzing data from an anonymous web-based survey conducted from March 26, 2020 to April 20, 2020 on 3259 individuals in France, Detoc *et al.* [40] reported vaccine hesitancy as the most significant obstacle in COVID-19 vaccination. Results suggested that about 77.6% of individuals were certain or probable to take COVID-19 vaccine and 47.6% people were inclined (certainly or probably) towards a clinical trial.

A number of studies conducted cross-sectional survey on vaccine hesitancy in Hong Kong and China. Wang *et al.* [41] conducted a study on 806 nurses in Hong Kong, China and found a low level of COVID-19 vaccine intent and high level of vaccine hesitancy. Wang *et al.* [42] studied on 2047 working individuals in Hong Kong, China about the change of intentions towards COVID-19 vaccines. They found a reduction in the willingness to accept COVID-19 vaccination in the third wave(34.8%) than the first wave(44.2%) among the survey participants. Researchers in [43] conducted a study to observe the patterns of vaccine hesitancy among different socioeconomic groups and residency status in Shanghai, China. The survey was conducted on 1021 parents of children aged below 18 years through a questionnaire survey. The result show that about 73.8%, 63.9%, and 52.4% of respondents expressed strong concerns on adverse side effects, safety, and effectiveness of vaccines, respectively. Additionally, they noticed that rural non-local residents of Shanghai showed more concerned about side effects, safety, and effectiveness of vaccine compared to local residents of Shanghai.

D. COVID-19 VACCINE ACCEPTANCE

A few studies investigated the vaccine acceptance level in US [44], [45] and abroad [3], [23], [46]. Majority of the studies underlined that people who are very concerned about COVID-19 infection and health care workers are more likely to accept the COVID-19 vaccine.

Kreps *et al.* [44] conducted a study to evaluate the factors that are associated with COVID-19 vaccination in the US.

Creating two hypothetical vaccine scenarios (Vaccine A and Vaccine B) based on vaccine efficiency, protection duration, major and minor side effects, approval from the US Food and Drug Administration (FDA), the origin of vaccine, and endorsement, the study explored wiliness of 1971 individuals to receive the vaccine by conducting a survey on July 9, 2020. The survey results show that about 56% of respondents are willing to accept the COVID-19 vaccine. An increase in vaccine efficiency and protection duration motivates people to accept the COVID-19 vaccine.

Goldman *et al.* [45] conducted a study in six countries (i.e., the US, Canada, Israel, Japan, Spain, and Switzerland) to understand the willingness of the parents having children aged less than 18 to accept a COVID-19 vaccine. Data were collected by administering a cross-sectional survey to 2557 respondents from March 26 to June 30, 2020. Survey results indicate that about 43% of respondents are willing to accept expedite testing and approval of the COVID-19 vaccine to make it available for use. Moreover, about 18.4% of respondents would allow their children to participate in the vaccine trial process. This study suggested that public health and vaccine providers should consult and understand the attitudes of the people to increase acceptance and administration of the vaccine.

An online-survey-based study on 2006 people (18+) [47] in the US indicated that the COVID-19 vaccine acceptability is around 69% among the respondents. Several factors, such as recommendation of healthcare provider, moderate or liberal leniency towards politics, perceived future likelihood of COVID-19 infection, effectiveness of vaccines have been identified as the key indicators whether people are willing to get vaccinated.

Upon conducting a review, Troiano and Nardi [23] evaluated the level of COVID-19 vaccine acceptance by the people worldwide and analyzed the factors that influence the tendency to accept COVID-19. Analyzing 15 published papers, this study found up to 86.1% acceptance of COVID-19 vaccine by students and 54.3% to 77.6% acceptance by the general people.

Bell *et al.* [3] conducted an online-based cross-sectional survey and semi-structured interviews to assess the acceptability of the COVID-19 vaccine of 1252 parents or guardians aged greater than 15 years who live in England with a child aged 18 months or less. The survey results show that about 55.8% and 48.2% of respondents are more likely to accept the COVID-19 vaccine for themselves and their children, respectively. Some of the respondents (i.e., 34.3% for themselves and 40.9% for children) were unsure but seemed to accept the COVID-19 vaccine. Thus, most people are interested to take the vaccine for self-protection from COVID-19.

Leng *et al.* [46] conducted a study in China to determine the preference of the people for the COVID-19 vaccine acceptance. A total of 1888 responses were collected from six Chinese provinces that were selected using a stratified random sampling method considering GDP per capita and geographical location. The survey reveals that trust in the

vaccine and vaccination process is considered as an important factor.

In summary, from the social media and cross-sectional survey it is evident that around 40% Americans are COVID-19 vaccine hesitant and a couple of factors are involved with this hesitancy. Some of the identified factors are: lack of trust in vaccine development, rush in vaccine development, politics, side effects of vaccine, and lack of knowledge about vaccine. It is also observed that female, younger generation, Black/Hispanic Americans, low education attainment, unemployment, not having a vaccine for influenza-like illness, anti-vaccine attitudes, religiousness and negatively associated with willingness to accept vaccine. On the contrary, individuals concerned about the COVID-19 infection and disease and health care workers are positively associated with the vaccine acceptance.

III. THE POWER OF PUBLIC OPINION

Public opinion is a powerful force and has been considered to shape national and international policies. In 1856, Abraham Lincoln famously said *"Our government rests in public opinion"* [48]. So also, others have emphasized the importance of public opinion in policy formation at various levels of governance [49], and we extend this principle to our study of public sentiment and opinion themes toward the COVID-19 vaccine. During the analysis, we found various interesting Tweets expressing positive and negative sentiment about COVID-19 vaccination, many of which reflect general tweet themes and are worth highlighting. For example, we observe a theme of gratitude, and some tweets are very thankful to science and the vaccination process: *"Just found out I get my 1st vaccine through my work as a Professor today and Totally unexpected, but very thankful for the science and medicine"; "Received my first dose of the COVID-19 vaccine today. Have to say I was impressed. They have 1500 scheduled today. It was a very smooth process. Thank you"; and "Getting my second dose of the Pfizer vaccine tomorrow. Grateful for the common grace God affords us through science. #vaccine #grace"*.

People also expressed thankfulness to the staff members working with vaccine delivery, as a dominant sub-theme of gratitude in the tweets: *"Got my first dose of the vaccine. Very appreciative of the staff here at Gateway. You are the real super hero"*. Administering vaccines provide people with a sense of safety and supportive staff and delivery processes instill a strong sense of comfort: *"I just wonder how safe I should feel after the vaccine and on the side if the world (and which parts) is really ready to accept travelers"; "Got the vaccine! So thankful for God's protection all around me!"*. In contrast, mixed emotions were displayed regarding the perceptions of vaccination privileges in certain states: *"More equity in vaccine rollout please"*. People were also concerned about the eligibility parameters in the early stages: *"are in home care workers (helping seniors with ADL) eligible for Covid19 vaccine?"*.

A fair proportion of Tweets have moderate to severe

negative sentiment, and negative sentiment increased from early February to late March of 2021. One of the significant reasons driving negative sentiment was the perceived side effects caused by vaccination: *"Sure happened to me. First vaccine Friday and I'm flat on my back with every side effect but the anaphylactic shock. I'm miserable."; "I wish the vaccine didn't inject so much bone hurt juice in my arm"; "I got the J&J vaccine; for me the side-effects have been extreme tiredness weirdly combined with difficulty falling asleep, chills and sweats, a couple of brief dizzy spells.; A bit rough but hey, it's clearly doing *something*"*. Some tweets were more specific complaints and some were caustic and reflected deep negative sentiment with probable preconceived bias: *"My son has a severe shellfish allergy. I have Arachnoiditis... patients r reporting bad side effects from the non FDA approved COVID-19 Vaccine. I'm unable to travel so don't need a Vaccine Passport..The government might as well tattoo us w a number. Sounds similar to the Nazis."*

Humor, as a theme, has not been lacking in its various forms such as subtle, sarcastic and obvious humor: *"As with almost everything else in my life the only side effect of the vaccine so far is a lot of gas"*. Other tweets appealed to scientific and logic concerns, including post-vaccination challenges: *"Vaccine prevents you from serious illness, the need to be hospitalized and death. You can still become infected and potentially spread to unvaccinated people who can then get very ill and die. The more who get vaccinated, less people will get sick and die. It's not a conspiracy."; "There is still too much discrepancies with CoVid19 vaccine. Have you heard about President Biden falling? Is he losing his coordination from his injections?"*. Rapid development of the vaccine undermined the trust for some people. Besides these, people also have negative comments about the rumoured vaccine passport in the US - *"If Vaccine Passports become reality, it will be necessary to have a nationwide boycott to end this unconstitutional madness"*. Other concerns regarding the COVID-19 vaccine include the development of potential vulnerabilities towards other diseases and reduced natural immunity: *"But do the vaccines have a tendency to trigger antibody dependent enhancement instead of preventing viral infection? If the vaccine does not result in a robust response in neutralizing antibodies, you might be at risk for more severe lung disease if you're infected with a virus!"*.

IV. PSS: AN APPLIED THEORETICAL FRAMEWORK

The "Public Sentiment Scenarios" (PSS) framework was initially developed to analyze COVID-19 public sentiment towards reopening the economy for the purposes of catalyzing and augmenting public policy formation [5]. Samuel *et al.* highlight the tremendous value of latent public sentiment and the need to capture it effectively for the purposes of influencing and contributing to the formation of public policies at all relevant levels, including federal, state and local governance. They carefully articulate the power of public sentiment, public perceptions and public opinions and posit that the collective wisdom of mass public spontaneous and

continuous-stream social media posts are reflective of such public sentiments, perceptions and opinions. They also reference a wide range of research which demonstrates the use of social media analytics to gauge public opinion, and the power of public sentiment towards a broad range of issues ranging from the stock market, to judiciary processes, political agendas, technology, crisis management policies and the study of human behavior [50]–[54]. Social media information is powerful and can pose contagion challenges through virality of posts, as the substance and format of information have been shown to impact human sentiment and performance [55], [56]. In the original PSS framework, the scenarios were presented as alternatives based on aggregate public sentiment conditions, and we adapt the PSS theoretical framework and apply that concept to simultaneous and emerging scenarios. We use the PSS framework in this study to structure our hypotheses for statistical analysis, and then to evaluate the public policy implications of the sentiment statistical analysis we perform on vaccine tweets.

SCENARIO ANALYSIS: 3 SIGNIFICANT SCENARIOS

The US is now (early May, 2021) in an enviable position globally, where on any given day, it has cumulatively more vaccines available than people willing to receive the same. News reports have indicated increased public hesitancy towards the vaccine partly due to the controversial federal government's decision to implement a 10 day Johnson & Johnson vaccine pause, and including those who have already taken the first dose but do not appear to be returning for the second vaccine dose in a timely manner. We use the PSS framework to analyze the dominant sentiment in February and March of 2021, and the two significant public sentiment changes - change in positive sentiment from early February to late March of 2021, and change in negative sentiment from early February to late March of 2021. We do not dwell upon neutral sentiment in our analysis. The two changes do not automatically correlate negatively as may be intuitively perceived, because it is possible that both positive and negative sentiment counts or scores may increase or decrease together at the expense of neutral sentiment:

- a) **Scenario 1: Positive public sentiment is dominant in February and in March, 2021.** Given this scenario, selected on the basis of descriptive analytics, we observe strong positive sentiment towards the vaccine. Illustrative tweets include celebratory posts such as: "@PurelyNumbers Yay for vaccine!..."; "@sprizee Happy birthday and congrats on the vaccine!"; "2nd dose ready. #vaccine #Happy"; "got a pfizer vaccine lined up for 4:20p on the afternoon before my birthday, nice nice nice nice"; "I love getting on Twitter / Instagram and seeing another friend, loved, or/and mutual receiving their vaccine. Just makes me so insanely happy" and "Got the vaccine! So thankful for God's protection all around me!". These posts demonstrate positive public perceptions based on high levels of confidence that vast sections of the population have in the vaccines.

There is also a strong positive public perception on the operations and delivery of the vaccine: "Super shout out to everyone at the Moorestown Mall COVID-19 vaccine mega site. Unbelievably efficient & well run....really impressive..."; "Today I got my second dose of the Pfizer COVID-19 vaccine. Kudos to the Vermont National Guard for a job well done: well organized, smooth, courteous."; "Best 'day-date' ever with @john_r_ratliff Our 1st vaccine dose is DONE. Thank you @UofL-Health for the massive undertaking, exceptional organization, and seamless process. #pfizercovid19vaccine #SeeYouIn3Weeks" and "Got my first COVID vaccine this morning administered by @NashvilleHealth. They are doing a phenomenal job with the logistics of rolling this out to our city. Everyone was friendly, helpful and efficient. Big thanks and kudos to all involved!!". These posts demonstrate public sentiment being driven by the vaccines and by smooth delivery of the same - an important takeaway from this for public policy is the emphasis on facilitating a comfortable and efficient vaccination process.

- b) **Scenario 2: Positive public sentiment decreases from early February to late March, 2021.** Interestingly, we saw a decrease in positive public sentiment as indicated by the change in the proportion of positive tweets from early Feb to late March of 2021. We do not speculate reasons for the same, but attempt to identify and discuss mass perceptions to help the public policy discourse. Skepticism, reasonable and unreasonable, appears to have increased which may have toned some proportion of positive tweets into neutral or negative territory: "@Ironman_E @cruddydre_ @Ms_Jaydee But if your vaccinated and the vaccine works, what does it matter if im vaccinated? Youre safe right? Because the vaccine works right?"; "If we're supposed to get the vaccine to protect other people but other people are supposed to get the vaccine to protect other people, who is it actually protecting? If the vaccine is already protecting those other people how does other people getting it further protect them?"; "excited to announce that i am not 30, therefore ineligible to receive a vaccine in the state of New York" and "@7_bunnies @TwixxBar07 @courtneymilan "Experts do not know how long" I'm protected by the antibodies I've acquired from infection. That's fine. They don't know how long the vaccine will protect, either, so we'll all learn together. Meanwhile, I'll stick to my own immune system's protection." . This change in sentiment, though minor and may be mitigated by a number of factors such as a general loss of enthusiasm without loss of confidence in the vaccine. Public policy can be quite useful in addressing such scenarios to prevent the erosion of positive public perceptions of vaccine and the vaccination process - this can be achieved through appropriate communications and pronounced narrative development.

c) **Scenario 3: Negative public sentiment increases from early February to late March, 2021.** Multiple sentiment scoring methods affirmed the proportional increase in negative sentiment from early February to late March of 2021. This is more alarming than the minor loss of positive sentiment across the same periods: a loss of positive sentiment can be attributed to a number of factors and not necessarily reflect a loss of confidence in vaccines, however a proportional increase in negative sentiment over time could imply eroding confidence in vaccines based on rumors or inaccurate information. There could also be statistically disproportionate emphasis on some cases of people falling ill after taking the first or both dose/s of the vaccines, or outlier events such as deaths following vaccinations. Illustrative tweets convey these concerns and negative sentiment quite effectively: *"Whoa...when they say that the second vaccine might have strange side effects, the are NOT kidding."*; *"Vaccine update: a day after shot one of the Pfizer vaccine I'm definitely feeling the arm pain, a headache, and kind of run down."*; *"@onewiththesand Oh no. Feel better. Do you feel it is from your vaccine shot?"* and *"I would not have taken the vaccine if I knew it would make me sick..."*. While there is a general concern about some uneasiness, discomfort and pain post-vaccination, most people appear to accommodate that, as illustrated by many tweets along the lines of: *"pretty achy and fatigued following my second vaccine dose yesterday. but still so grateful and hopeful. worth it 1000x over! #GetVaccinated"*. Perhaps, the greater challenge is posed by some factual news on statistical outlier cases augmented by rumors of people being severely affected by vaccinations or even dying from the vaccine. This is illustrated by the deeply negative sentiment in some of the vaccine tweets in March of 2021: *"@cruddydre_ My uncle got the vaccine and died a week or 2 later. I'm good bro yal can have that vaccine."*; *"@DanMac2014 @Kayla_Grey @pfizer Wish one could request your vaccine, @moderna was a monster. Haveing had COVID recovered in 3 days with light fever/aches. Vaccine-1st dose F'd me up-5x worse then VID did. I wanted to crawl under a rock & die the past 3 days. Not sure I'll go for 2nd..."*; *"Mom don't be a cry baby. It doesn't help that I have a sister that tried to have her not take it telling her that there's people who died from the vaccine...I tell my mom that my sister isn't well informed...directs mother to the daughter who has a PhD."*; *"@princessnofrog Haven't seen that yet they are not sure because it's still in the experimental stage but there has been people that have died from the vaccine. Meaning it can be deadly" and "Maybe the vaccine will kill me since the virus didn't"*. Such expressions of fear, despair and negativity tend to spread fast through social media and personal messaging and communications. Public policy is critical in such scenarios and proactive steps are required to

build public confidence. Transparency on adverse events must not be compromised - post vaccines illnesses and the rare cases of deaths must continue to be documented and such information must be made publicly available. However, public policy must mandate resources for such information to be augmented and qualified by factual information truthfully highlighting the details along at least two lines of reasoning: a) An emphasis on the rarity of extreme events such as vaccination induced deaths and b) disassociation of vaccination as the cause of deaths after vaccinations, where applicable, and clarity on the identified or potential causes of death.

V. METHODS

Thus far, we have used a broad range of secondary data on COVID-19, and on vaccines, along with analysis of extant research to provide a descriptive narrative of the complex scenarios surrounding the vaccination and herd immunity challenge. We have also drawn some descriptive insights from our primary dataset which consists of tweets from Spring 2021, to provide a better understanding of the scope and direction of public opinion.

A. DATA ACQUISITION AND PREPARATION

We used the `rtweet` package in R software to download publicly visible tweets in early February and late March of 2021, using a keyword filter *"vaccine"*. After cleaning and preprocessing the downloaded data consisting of over a million tweets, and excluding spam by `url` posts, we subset 5,131 and 9,036 tweets filtered by country as *"United States"* for Feb'21 and March'21, respectively. In addition to cleaning the data to reduce special characters, we also ran a custom algorithm to replace most of the identifiable abusive words in tweets with a unique character string *"7abuvs12304"* with a low likelihood of natural occurrence in tweets. We applied the `SentimentR` package, with the `Jockers` dictionary, to compute the sentiment score for each tweet [57]. The tweets were then classified into three sentiment classes on the basis of tweet sentiment scores: positive sentiment (score > 0.10), negative sentiment (score < -0.10) and neutral sentiment ($-0.10 \leq \text{score} \leq 0.10$). Various characteristics of the tweets were explored using descriptive analytics, and tweet lengths were seen to have no significant correlation to sentiment scores (Fig. 2). Tweet lengths and sentiment scores displayed bimodal and skewed visually undetermined distributions. The majority of tweets are between 10 and 55 words in length, and most of the tweets sentiment scores fall in the range of -1 to 1.

Fig. 4 shows day-wise boxplots representing the distribution of sentiment scores in three different classes (positive, negative, and neutral) based on the Feb'21 dataset (Fig. 4a), and March'21 dataset (Fig. 4b). From both of the data visualizations, it can be observed that the median values of the positive sentiment scores are equal or slightly greater than 0.25, whereas the maximum values are usually greater than 0.50. on the other hand, the median values for the negative

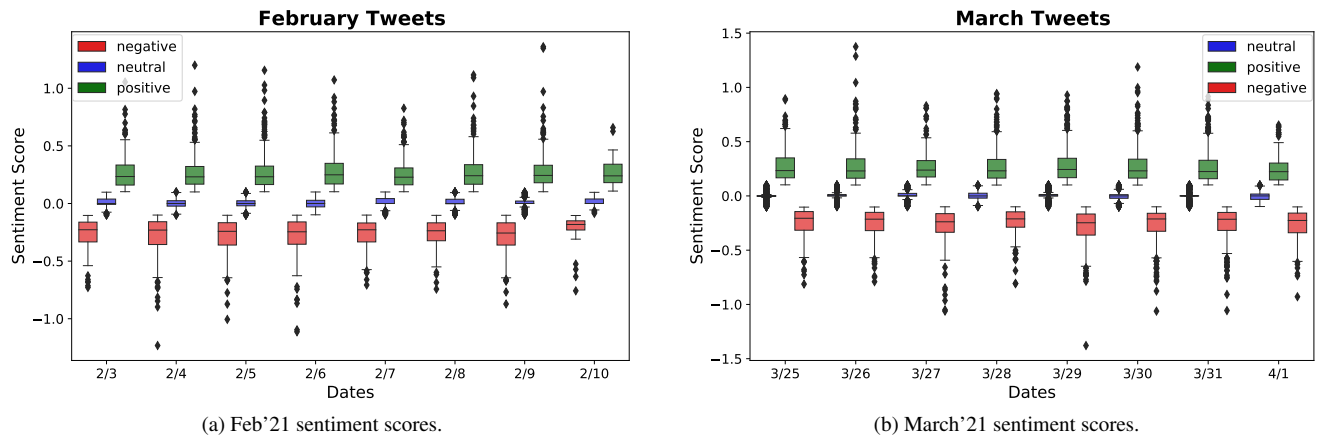


FIGURE 4: Boxplots for the sentiment scores.

sentiment scores are close to -0.25, and minimum values are usually smaller than -0.50. For both positive and negative sentiments, the median value is not located in the middle of IQR (interquartile range) of the boxplot, which implies that the sentiment scores are most likely not normally distributed. Additionally, we observe more outliers in positive sentiment boxplots than in negative sentiment boxplots (more tweets in extremely positive class than extremely negative). Descriptively, these boxplots indicate that the overall strength of positive sentiment is greater than that of the negative sentiment.

B. COMPUTING AND COMPARING SENTIMENT SCORE

Our analysis of the sentiment scores shows that the percentage of tweets with positive sentiment is higher (daily as well as overall, in both datasets) compared to the percentage of tweets with negative sentiment (Fig. 3). As mentioned earlier, by default, for computing the sentiment of each tweet we have used the R package `sentimentR` along with the default dictionary 'Jocker' [57]. To further validate the sentiment score calculations, we also used two more lexicon-based Python libraries `Textblob` and `VADER` (valence aware dictionary for sentiment reasoning). `Textblob` provides a simple API for text mining, text analysis and text processing. It reuses the corpora of the Natural Language Toolkit (NLTK) for text analysis. `VADER` is another lexicon and rule-based python library, specially designed for analyzing sentiment of social media text [58]. Oyeboode and Orji in [59] show that `VADER` performs significantly better in terms of sentiment scoring accuracy than `Textblob` when analyzing social media text. Fig. 5a and 5b show the comparison of three methods to identify sentiment classes such as positive, negative and neutral on the datasets from Feb'21 and March'21, respectively. It can be observed that all the methods have a similar trends in classifying tweets. Each method identifies a higher percentage of positive tweets compared to the negative tweets with a significant percentage of neutral tweets. Considering the superior accuracy of `VADER` over `Textblob` and the similarity of results with `sentimentR`

to identify positive and negative classes, we have used the sentiment score obtained by applying `sentimentR` in the rest of our analysis.

C. WORD AND PHRASE ASSOCIATIONS

The text corpus created from the filtered and cleaned tweets was used for word and phrase analysis to gain insights into dominant themes. We have used word frequency and N-grams analysis to study the text corpus as shown in Fig. 6 (Fig. 6a with crossed pattern for Feb'21 tweets and Fig. 6b with dotted pattern for March'21 tweets). Word frequency analysis provides an array of high frequency words, sorted by decreasing frequency: "vaccine, covid, pharmacy, technician, people, health, job". Many of the high frequency words were observed to be part of tweets supporting positive sentiment about vaccination. N-grams identify frequently used word pairs and word sequences, and they reveal interesting patterns. The top Bigrams (two word sequences) included: "covid vaccine, pharmacy technician, get vaccine, vaccine support, technician covid, cvs health", and the Bigrams also indicated frequent usage and support for positive sentiment about vaccination. The top Trigrams (three word sequences) were more insightful about the mass focus on Vaccine distribution and operations, and the top items included: *covid vaccination support, pharmacy technician covid, technician covid vaccine, technician retail store, support cvs health*, which were mostly aligned with support for positive sentiment about vaccination. The top Quadgrams (four word sequences) included: *pharmacy technician covid vaccine, technician covid vaccine support, pharmacy technician retail store, vaccine support cvs health* and similar phrases indicating discussions about CVS, jobs and vaccine support. The Quadgrams also show reasonable support for positive sentiment about vaccines, and emphasized the operations and delivery aspects of the vaccine. Lower frequency Quadgrams provided insights into emerging themes or common concerns, including concerns about the governments arranging for sufficient vaccine quantities and challenges.

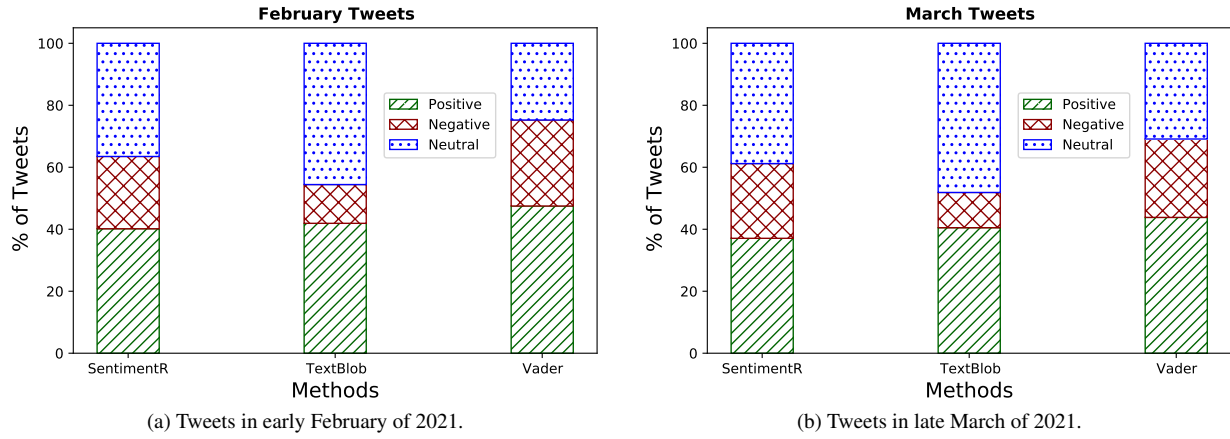


FIGURE 5: Sentiment classification using three different methods.

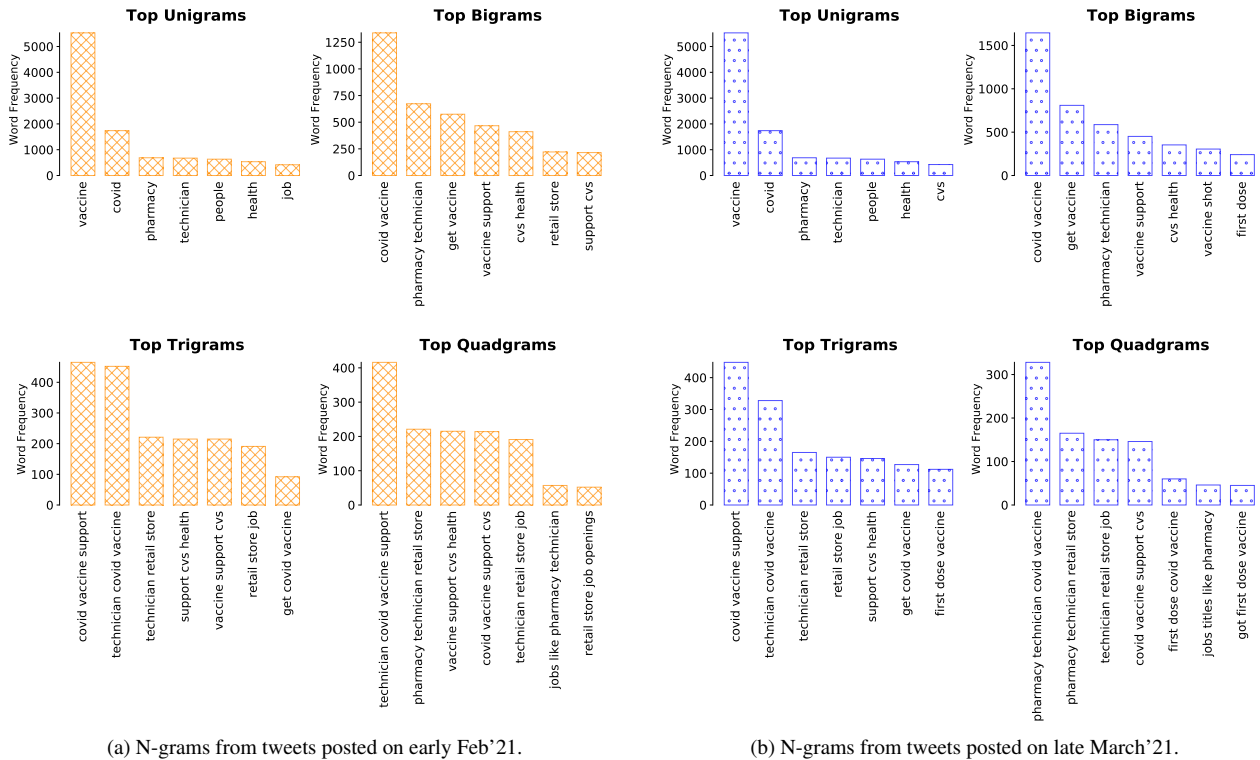


FIGURE 6: N-grams.

D. GEO-TAGGED ANALYTICS

The state name of each tweet is determined based on the tagged geo-locations of tweets within the contiguous United States. Among 5131 tweets in early February, 2021 we have a total of 5095 geolocated tweets. Also, among 9036 tweets in late March, 2021 we have a total of 8961 geolocated tweets in the data. The *Syuzhet* package in R is utilized to calculate the scores of eight sentiments including anticipation, joy, trust surprise, fear, sadness, anger, and trust. The aggregated sum of each sentiment score is calculated for each state, then normalized by the number of tweets of each state. The percentages of positive, neutral, and negative sentiments are

calculated at the state level as well. The state-wise sentiment is discussed in Section VI-B.

E. TEMPORAL DATA ANALYTICS

COVID-19 vaccine sentiments among the Twitter users are observed and analyzed over a period of 6 days, both in early Feb'21 and late March'21 time frames. In the Feb'21 time frame, an increasing positive sentiment trend, a decreasing negative sentiment trend, and unchanged neutral sentiment trend are found (Fig. 3a). On the other hand, in the March'21 time frame, at the end of the week, there is no significant changed is observed in the positive and neutral sentiment

trends, but a slight increase is observed in the negative sentiment trend (Fig. 3b). Apart from this, it is also noticed that the negative sentiment score is the lowest on Monday for both the time frames. Negative scores are higher over the weekends compared to the weekdays. On the contrary, the neutral sentiment score is the lowest during the weekends. It suggests that people post more negative tweets in the weekends than weekdays.

Overall, from Fig. 7, it is noticed that positive sentiment score over time decreases (from early Feb'21 to late March'21). On the other hand, although the negative sentiment does not change much, the neutral sentiment increases noticeably. This indicates that while COVID-19 vaccination was administered in the US with full swing in February and March of 2021, the perceptions about vaccination did not improve in the US.

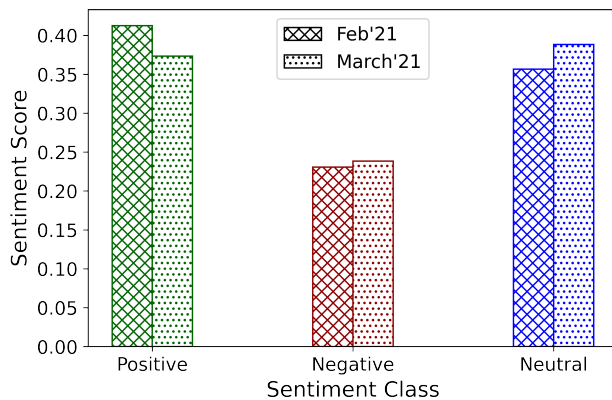


FIGURE 7: The change of sentiment score from early February to late March of 2021.

To understand the effect of type of Twitter users on the sentiment score, all the Twitter accounts are segregated into two groups based on the account created year. Those who have created their Twitter account before January 2020, are categorized as mature group and those who created after January 2020 are categorized as new group. The tweet sentiments based on these age groups are observed and plotted in Fig. 8. Out of the total tweets, a total of 6% and 6.21% tweets are posted by new users in early February and late March of 2021, respectively. Results indicate that the new user tweets did not dominate the overall COVID-19 vaccine sentiments. Among the Feb'21 tweets, about 40% are mature positive, 21.15% are mature negative, and 33.15% are mature neutral tweets. On the other hand, among the March'21 tweets, 35.30% are mature positive, 22.09% are mature negative and 36.40% are mature neutral tweets. It is observed that, both in Feb'21 and March'21 tweets, positive sentiment is higher than negative sentiment among the mature users. Additionally, percentage of neutral sentiment is increased among the mature users from February to March 2021. However, no such significant change is observed for the new users.

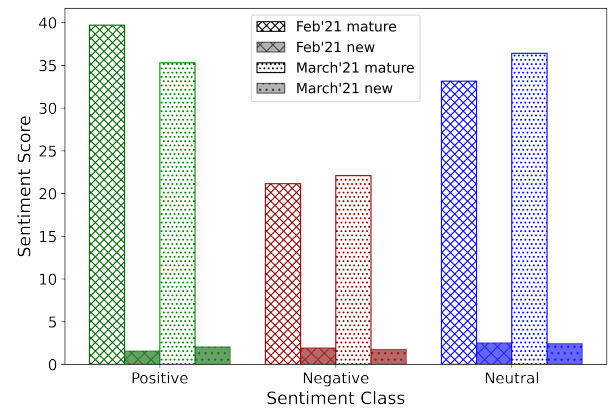


FIGURE 8: Temporal sentiment analysis in early February and late March of 2021 by the type of Twitter account users (new users vs. mature users).

VI. RESULTS AND DISCUSSION

In this section, we are going to provide the statistical validation of our computed sentiment values, followed by the US state-wise sentiment analysis and then US state-wise actual vaccination scenarios analysis.

A. STATISTICAL ANALYSIS OF SENTIMENT VALUES

The descriptive analytics presented thus far provides reasonable support for a dominant positive sentiment on vaccines than the corresponding negative sentiment for both February and March of 2021. So also, there exists some initial evidence for a decrease in positive sentiment from Feb'21 to March'21, and increase in negative sentiment from Feb'21 to March'21. This is also evident from the visual analysis of sorted sentiment scores for early February and late March of 2021, as shown in Figs. 9a and 9b. However, this is insufficient from a statistical perspective and it is necessary to validate such descriptive findings. We validate the exploratory findings with a Proportion test on the ratio of positive tweets to negative tweets separately for early February and for late March, and further verify this by applying an Exact Binomial test as shown in Tables 1 ~ 3.

1) Dominant Positive Sentiment for February and March of 2021

We employed the Shapiro-Wilk normality test for the sentiment scores for Feb'21 (Table 1) and for March'21 (Table 2), and the tests demonstrated that the sentiment scores were not normally distributed. The results were also supported by the Q-Q plots indicating departure from normality (Fig. 10). We therefore employed the same validation methods used for the PSS framework by Samuel *et al.* and applied the non-parametric Proportion Test and Exact Binomial test for the analysis [5]. The results of the Proportion Test, with the null hypotheses indicating that the proportional count of negative tweets is equal to or greater than positive tweets, indicated that such a null hypothesis could be safely rejected with a P-

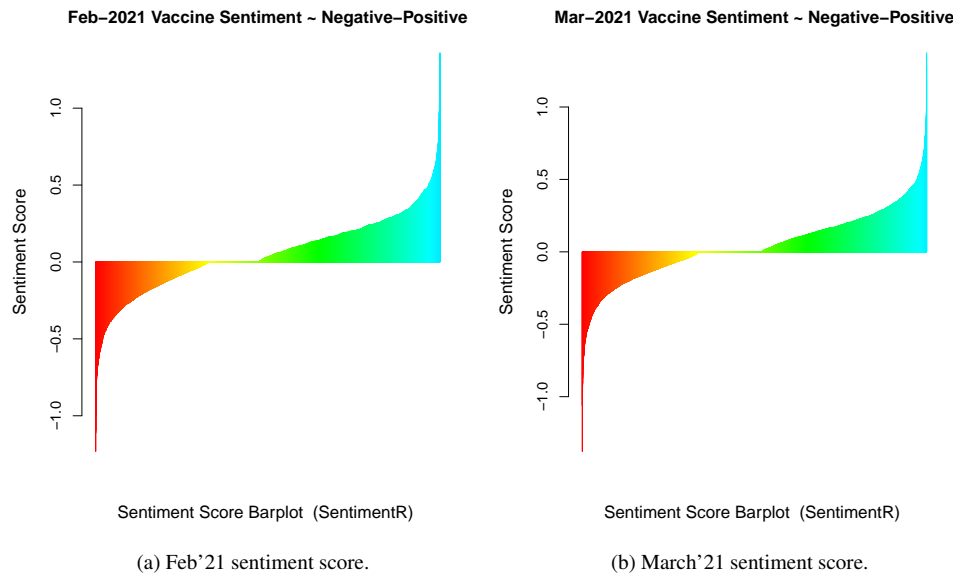


FIGURE 9: Positive and negative sentiment score.

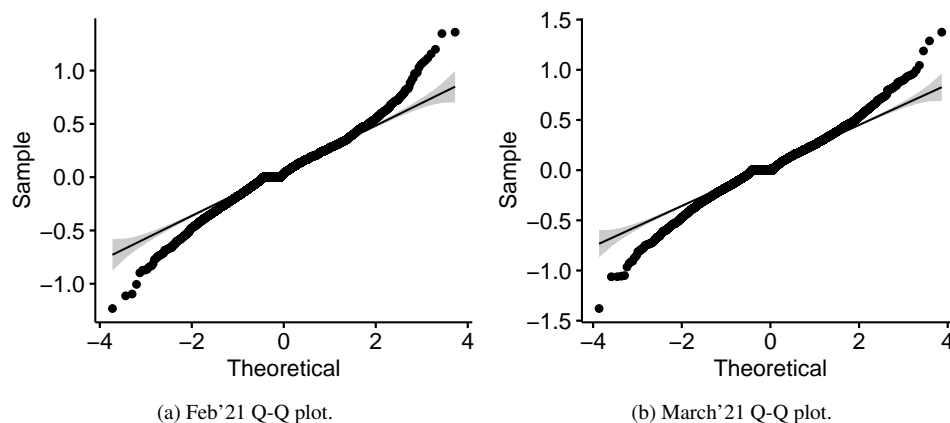


FIGURE 10: Q-Q plots.

Value significantly <0.0001 . The Proportion Test results thus validated the alternative hypothesis for the positive sentiment proportion being significantly greater than corresponding negative sentiment. The Exact Binomial tests confirmed the same results. Finally, the Wilcoxon signed rank test was used to confirm the findings, as the sentiment scores were not normally distributed as indicated by the Q-Q plot in Figs. 10a and 10b. Thus, the dominance of the positive sentiment scores for both February and March were supported both by the Proportion Test, as well as by the Exact Binomial Test.

2) February-March: Lower Positivity and higher Negativity

Though, as demonstrated above, positive sentiment for both February and March remained stronger than corresponding negative sentiment, yet our analysis of relative change highlighted important insights: Relative proportion of positive

sentiment decreased from February to March, and relative proportion of negative sentiment increased from February to March of 2021. As above, we employed the Proportion Test and further validated it using the Exact Binomial Test. Our null hypotheses was that positive sentiment remained at the same relative proportion or increased from February to March. The results as shown in Table 3 indicates that the null hypothesis can be fairly rejected in favor of the alternate hypotheses that relative proportion of positive sentiment decreased from February to March of 2021. So also, the relative proportion of negative sentiment increased from February to March of 2021.

B. STATE-WISE SENTIMENT ANALYSIS IN THE US

In this subsection, we analyze the US state-wise sentiment progression from early February to late March of 2021.

TABLE 1: Statistical tests for positive and negative sentiment counts for Feb'21.

(a) Proportion test for positive sentiment.

	Proportion Test	Exact Binomial test
Null Hypothesis	-ve senti \geq +ve senti	-ve senti \geq +ve senti
Alt. Hypothesis	True P is less than 0.5	True P is less than 0.5
P-Value	2.20E-16	2.20E-16
Conclusion	Reject Null Hypothesis	Reject Null Hypothesis

(b) Shapiro-Wilk normality test for sentiment scores.

	Shapiro-Wilk Normality Test
Null Hypothesis	Normally distributed data
Shapiro-Wilk statistic	0.98387
P-Value	2.20E-16
Conclusion	Reject Null that Data is normal

(c) Wilcoxon signed rank test for mean of sentiment scores.

	Wilcoxon signed rank test
Alt. Hypothesis:	True location is greater than 0
V	6117396
P-Value	2.2E-16
Conclusion	Reject Null Hypothesis

TABLE 2: Statistical tests for positive and negative sentiment counts for March'21.

(a) Proportion test for positive sentiment.

	Proportion Test	Exact Binomial test
Null Hypothesis	-ve senti \geq +ve senti	-ve senti \geq +ve senti
Alt. Hypothesis	True P is less than 0.5	True P is less than 0.5
P-Value	2.2E-16	2.2E-16
Conclusion	Reject Null Hypothesis	Reject Null Hypothesis

(b) Shapiro-Wilk normality test for sentiment scores.

	Shapiro-Wilk Normality Test
Null Hypothesis	Normally distributed data
Shapiro-Wilk statistic	0.98343
P-Value	2.20E-16
Conclusion	Reject Null that Data is normal

(c) Wilcoxon signed rank test for mean of sentiment scores.

	Wilcoxon signed rank test
Alt. Hypothesis:	True location is greater than 0
V	16629414
P-Value	2.20E-16
Conclusion	Reject Null Hypothesis

Fig. 11 illustrates the percentage of positive, negative, and neutral sentiment in different states in early February and late March of 2021. The positive sentiment is decreased in most of the states except a few states including New Mexico, Nebraska, Iowa, North Dakota and Alabama. The positive sentiment is increased by only 10% in these states. The largest decline in positive sentiment is observed in Vermont and Arkansas from more than 50% to below 20% and 20-30% respectively. A few states have around 20% decline in positive sentiment including New Hampshire, Michigan, South Dakota, and Tennessee, whereas South Dakota and Tennessee have significant increase in negative sentiment. Most of the states have none or very little change in negative

TABLE 3: Statistical tests for positive and negative sentiment counts for Feb'21 and March'21.

(a) Tests for change in positive sentiment from Feb to March.

	Proportion Test	Exact Binomial test
Null Hypothesis	Mar'21 +ve senti \geq Feb'21 +ve senti	Mar'21 +ve senti \geq Feb'21 +ve senti
Alt. Hypothesis	True P is less than 0.5	True P is less than 0.5
P-Value	0.03262	0.03262
Conclusion	Reject Null Hypothesis	Reject Null Hypothesis

(b) Tests for change in negative sentiment from Feb to March.

	Proportion Test	Exact Binomial test
Null Hypothesis	Feb'21 -ve senti \geq Mar'21 -ve senti	Feb'21 -ve senti \geq Mar'21 -ve senti
Alt. Hypothesis	True P is less than 0.5	True P is less than 0.5
P-Value	0.01159	0.01158
Conclusion	Reject Null Hypothesis	Reject Null Hypothesis

sentiments. A significant decline in negative sentiments is observed in only four states: Wyoming, North Dakota, Oregon, and Maine. A significant decline in both positive and negative sentiments in Arkansas results in a sharp rise of neutral sentiments. In contrast, both positive and negative sentiment are increased in a few states while neutral sentiment is decreased such as Iowa.

Fig. 12 shows the changes in two prominent sentiments, –fear and trust– between early February and late March in different states in the US. Fear sentiment against COVID-19 vaccine is very strong in most of the states, which is mostly increased over time. Near about two-third of the contiguous US states show overall fear score more than 0.4 in early April. In late March, 38 states have fear score higher than 0.4. Fear sentiment remain unchanged in the North-West and the West regions and increased in many states in the South-West, the South-East, and the Mid-West. It is observed that fear sentiments decreased or remain unchanged in big states including Texas, New York, California, Arizona, and Florida.

In contrast, trust sentiment is mostly decreased in majority of the states across the US except a few states including South Dakota, Nebraska, Oklahoma, Idaho, Arkansas, and North Dakota. Most of the states shows relatively moderate to low score in trust sentiments. Trust on vaccine is decreased significantly in Arkansas. Although, fear sentiment is unchanged in big states, trust sentiment is decreased in California, Texas, and New York.

C. VERIFICATION AND VALIDATION WITH ACTUAL STATE-WISE VACCINATIONS

Despite a recent decline in vaccine administration in the US (Fig. 1), the President of the United States set a goal on May 5, 2021 to achieve a 70% of vaccination of American adults by July 4, 2021 [60]. Thus, considering the importance of vaccination to protect public health, we explored the US state-wise actual rate of vaccinations after collecting vaccination data from the US CDC [12]. We observed that vaccination rate varies a lot from one state to another. Fig.

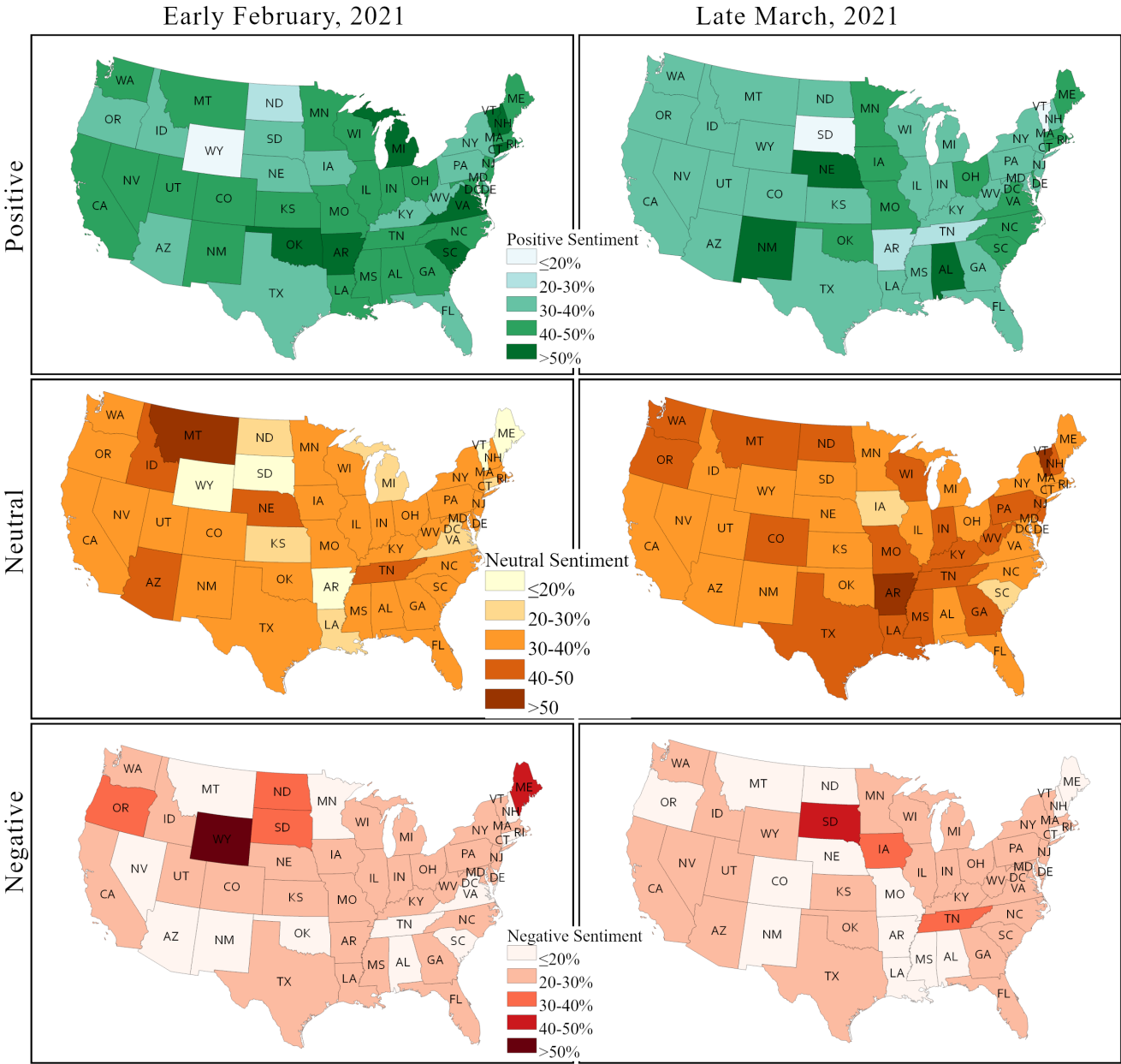


FIGURE 11: Changes in positive, neutral, and negative sentiment in different states in the US.

13 demonstrates the rate of fully and at least one dose of vaccination (%) in different states of the US, as of April 28. The figure indicates that people living in the Northeast regions of the US (e.g., Maine, New Hampshire, Connecticut, Rhode Island, Vermont, New Jersey, Massachusetts, New York, Pennsylvania, Delaware, and District of Columbia, and Maryland) have a higher rate of vaccine uptake. In contrast, a lower rate of vaccine uptake is observed among the people living in the Southern region of the US (e.g., Alabama, Georgia, Louisiana, Mississippi, Tennessee, Arkansas, Indiana, West Virginia, South Carolina, and Texas). However, a higher rate of vaccine distribution and administration is observed in Virginia.

A mixed feelings is noticed among the residents of the Midwest and West regions of the US (i.e., some states have higher vaccination rate, while some have lower). People living in the West region particularly in California, New Mexico, Arizona, and Colorado showed a higher willingness accompanied with a higher rate of vaccination. However, residents from Wyoming and Idaho showed a reluctance to accept vaccines, thus a lower rate of vaccination is observed in these two states. Although a higher use of distributed vaccine is observed in Utah, overall vaccine administration is low in this state. The States located in the Midwest regions also have a higher rate of vaccine uptake. More specifically, people living in South Dakota, North Dakota, Wisconsin,

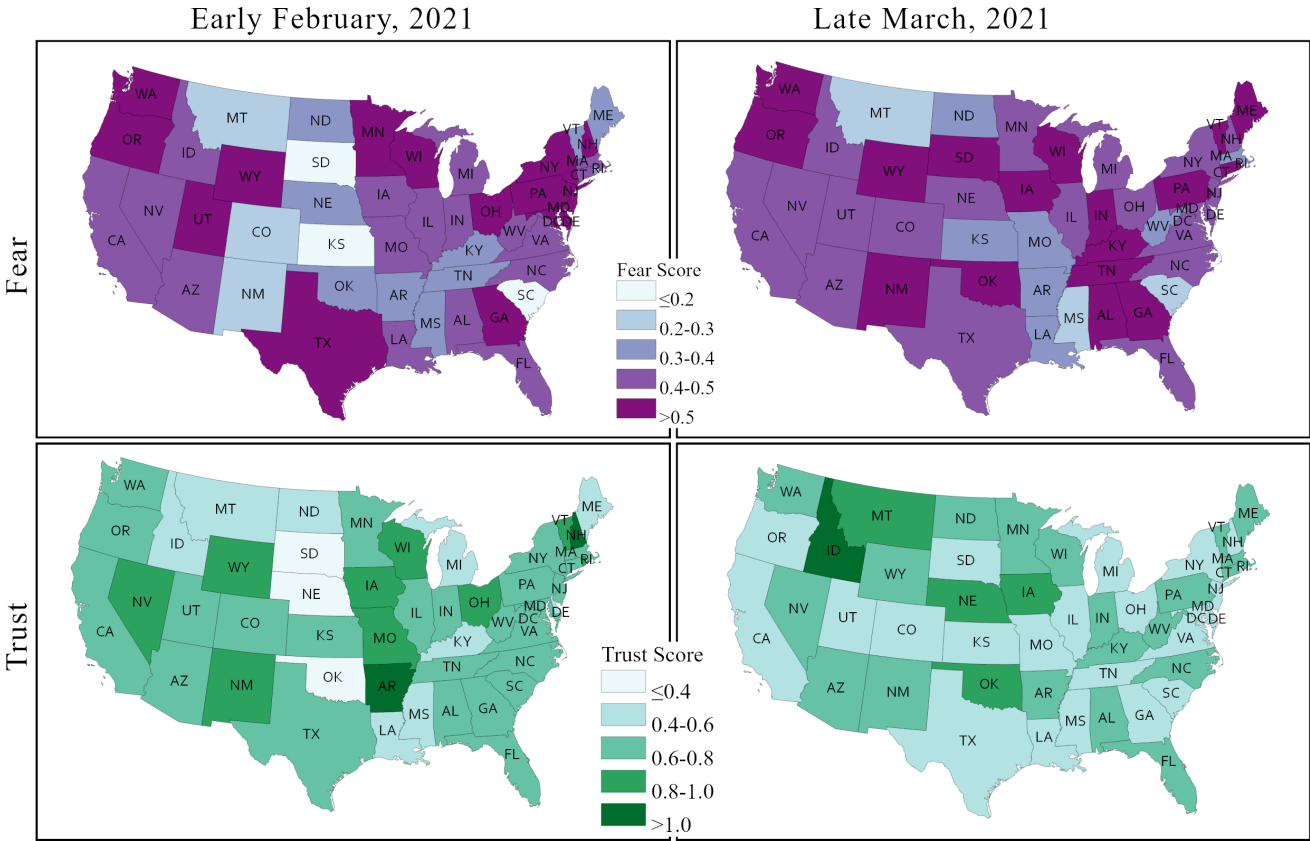


FIGURE 12: Changes in fear and trust sentiments in different states in the US.

Iowa, Kansas, and Minnesota have a higher rate of vaccination. On the other hand, Missouri and Indiana have a lower rate of overall vaccine administration.

To compare with our tweet sentiment analytics, we have also investigated the changes in vaccination uptake in the US upon collecting an average daily vaccination data from February 3 to 10 and average daily vaccination from March 25 to April 1 (Fig. 14). The figure explains that most of the states achieved an increase in vaccination in late March than early February. The highest increase in vaccination is observed in the Northeast regions of the US. Similar, a significant increase in vaccination is also observed in some states of Midwest regions (e.g., Kansas, Iowa, South Dakota), despite a low increase in vaccination in North Dakota and Minnesota. In contrast, the lowest increase in vaccination is noticed in parts of West region (e.g., Utah, Idaho, and Wyoming). However, California achieved a higher increase in vaccination in early February than late March. Similarly, many states in the South regions (e.g., Mississippi, Louisiana, Alabama, Oklahoma, Arkansas, and West Virginia) achieved a relatively low increase in vaccination in late March compared to early February. It is also observed that sparsely populated states and states with a lower number of population achieved a higher rate of vaccination in comparison to the highly populated states.

Extant studies mentioned that one of the main reasons for

a low acceptance of vaccines by the people is inadequate supply and distribution of vaccines and low accessibility to vaccines. Considering this issue, we have also investigated the share of vaccine dose used from the actual distribution in different states of the US to gauge the status of vaccine dose utilization, as of April 28 (Fig. 14). The figure shows a comparative scenario of differences between vaccine used and vaccine distributed across different states in the US. This figure verifies that states located in the Northeast regions have a higher percent use of the distributed vaccines. Similarly, a significant use of distributed vaccine is observed in states in the West and Midwest regions. In contrast, states in the South region showed a lower percent use of the distributed vaccine which verifies the lower rate of vaccination uptake in those states as shown in Fig. 13.

To understand the root causes of vaccine hesitancy among the Americans, a recent nationwide Household Pulse Survey (HPS) conducted by the US Census Bureau (from April 14 to April 26, 2021) investigated the status of COVID-19 vaccine hesitancy in the county-level of each state [13]. Respondents were asked “Once a vaccine to prevent COVID-19 is available to you, would you... get a vaccine?” with following four options: (1) “definitely get a vaccine”, (2) “probably get a vaccine”, (3) “probably not get a vaccine”, and (4) “definitely not get a vaccine”. The survey defined “probably not” and “definitely not” as the vaccine hesitancy, while “definitely

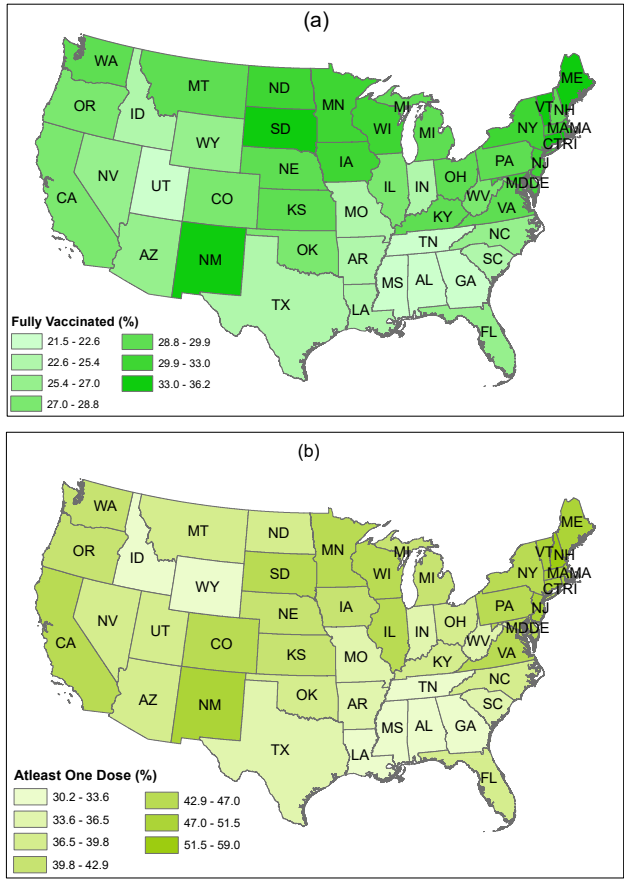


FIGURE 13: (a) People fully vaccinated per 100 persons (b) People with at least one dose of vaccine per 100 persons (Data source: [12]).

not" option is defined as the extreme vaccine hesitancy. Spatial representation of vaccine-hesitant in Fig. 15 indicates that many counties in the states of West region, part of Midwest region, and part of South region have a higher number of vaccine-hesitant and extreme vaccine-hesitant. Thus, a lower rate of vaccine used have been observed in most of these states.

Table 4 shows the prominent reasons of COVID-19 vaccine hesitancy in the HPS survey [13]. Side effects of vaccines is the main reason (20.86%), whereas, safety (17.04%) and trust (12.55%) are the other two important reasons, why Americans are hesitant to take the COVID-19 vaccine.

In another study, YouGov in association with the Institute of Global Health Innovation (IGHI) at Imperial College London (ICL) conducted surveys in 29 countries to gather global insights on people’s behaviors in response to COVID-19 [61]. Respondents, who did not receive a COVID-19 vaccine, were asked to provide their responses on a scale of 1 (i.e., strongly agree) to 5 (i.e., strongly disagree) to the question: "If a COVID-19 vaccine were made available to me this week, I would definitely get it". Only 34.5% respondents from the US expressed their willingness to receive a COVID-19 vaccine

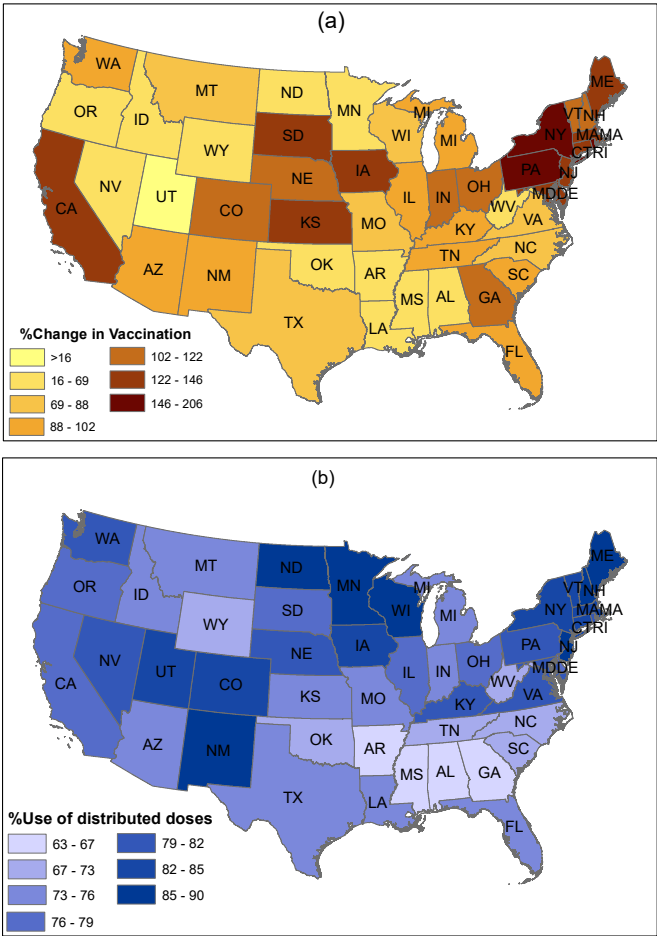


FIGURE 14: (a) Relative changes in vaccination between February (average vaccination from 3 to 10 February) and March (March 25 to April 1), 2021, (b) Percentage use of distributed vaccines, as of April 28 (Data source: [12]).

TABLE 4: Reasons not to receive a COVID-19 vaccine in the US (Data source: [13]).

Key Reasons to affect vaccine administration	Percentage
Concerned about possible side effects	20.86
Plan to wait and see if it is safe	17.04
Do not trust COVID-19 vaccines	12.55
Do not trust the government	9.63
Do not believe I need a vaccine	8.32
Do not know if a vaccine will work	8.32
Other people need it more right now	7.11
Do not like vaccines	4.39
Doctor has not recommended it	2.54
Concerned about the cost	1.42
Other	7.15
Did not report	0.68

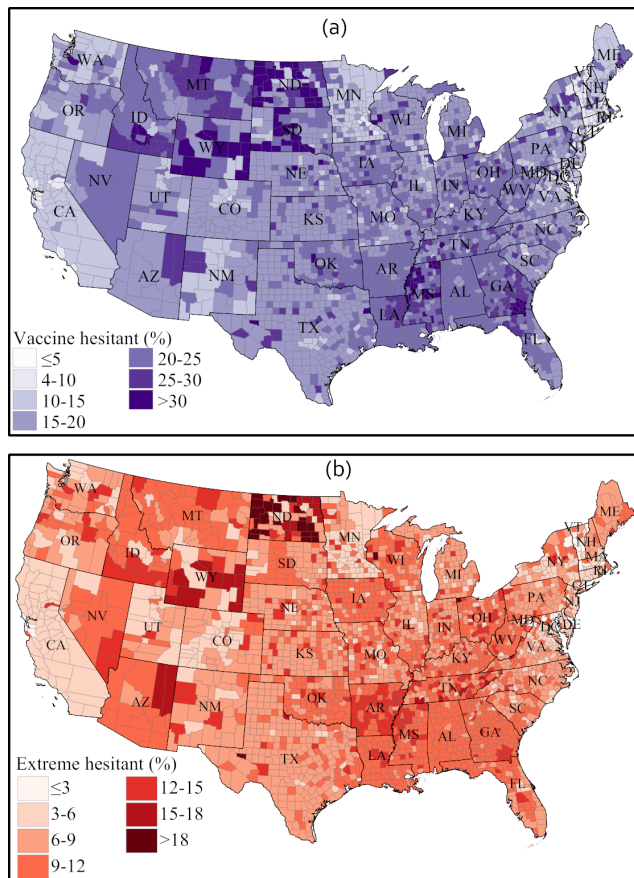


FIGURE 15: (a) COVID-19 vaccine hesitant (%) in the US (b) COVID-19 vaccine extreme-hesitant (%) in the US (Data source: [13]).

which validates our findings from Twitter sentiment analysis (i.e., roughly 37% of US Twitter users expressed positive sentiment on the COVID-19 vaccine (Fig. 7)). In that survey, UK has the maximum positive responses (67.2%), followed by Denmark (66.8%), Norway (58.5%), Germany (58%), and Italy (57.3%).

D. DISCUSSION

From our sentiment study (Figs. 11 and 12) and the US state-wise actual vaccination scenarios (Figs. 13 ~ 15), we found that both the public sentiment and vaccination administration scenarios vary from one state to another state and has a strong correlation with the population density and geographic location. Some of the notable key points are stated as below.

- 1) We observed that smaller states and sparsely populated states have a higher vaccination rate and a stronger positive vaccination sentiment. Illustratively, Maryland, Connecticut, Maine, New Jersey, Vermont, Delaware, South Dakota, North Dakota, Montana, Iowa, Minnesota etc. have more vaccinated proportions of people than some other states.
- 2) The Northeast regions have higher vaccination acceptance (e.g., Maine, New Hampshire, Connecticut,

Rhode Island, Massachusetts, Vermont etc.).

- 3) The maximum vaccine hesitancy is found in the southern regions (e.g., Alabama, Georgia, Mississippi, Tennessee, Arkansas etc.).
- 4) The Midwest and West regions show mixed vaccine acceptance scenarios. Notably, Wyoming and Idaho show lower vaccine acceptance. In contrast, California, New Mexico, Arizona, and Colorado show a higher vaccination acceptance. The HPS survey found that side effects, trust in vaccine and authority, safety, and efficiency are the main reasons for the lower vaccination acceptance in these states.
- 5) Although we found that in the sentiment progression from Feb'21 to March'21, positive sentiment declines and neutral sentiment increases in most of states in US, however, in terms of actual vaccination, we found that most of the states improved the vaccine administration rate. We unraveled that this is because of the increasing state-wise vaccine availability as the time progresses.
- 6) Among negative sentiments (fear, sadness, anger, disgust), fear is the strongest sentiment which is increased over time in most of the states.
- 7) Fear sentiment remain unchanged (between early February and late March) in the populated states like California, Texas, New York), whereas a slight decline in trust sentiment is observed in these states.
- 8) Changes in the sentiments are very abrupt among the small states (population) in the central parts of the US including Wyoming, South Dakota, Kansas, New Mexico, and Arkansas.

In summary, as vaccination hesitancy differs significantly from one to state to another, the state-wise policy implications also need to be different to increase the vaccination uptake. In the following section, we stipulate the possible public policies that might be implemented by local, state and federal government to make the sufficient population immune to COVID-19.

E. LIMITATIONS

The current study significantly contributes to the scientific community by investigating the public perceptions of the COVID-19 vaccine in the US amid numerous misinformation on vaccines, lack of trust in vaccine and authority, and vaccine hesitancy. Despite the timely contributions to the literature, this study also has some limitations which are stated below.

- First, although researchers and policymakers have used Twitter data widely, it inadequately represents perceptions and opinions of the people from different strata of the society [18], [62]. Many people (e.g., low-income and low educated people, elderly) do not have access to Twitter, thus, using Twitter data inaccurately characterizes the sentiment of the people towards vaccination.
- Second, Twitter data is susceptible to the bot activities [5]. Twitter bots are programmed to mimic daily Twitter

users and share contents of a specific subject matter to spread the news to the broader audience, which could be hoaxes and fakes and alter perceptions of the actual Twitter users [63]. Moreover, voluntary and involuntary mistakes by the users provide inaccurate data. Thus, the quality of Twitter data is reasonably affected by bot activities and user mistakes.

- Third, the sentiment scores extracted from Twitter using the *Syuzhet* package have some degree of inaccuracy for sentiment scores assigned to individual tweets, which may influence sentiment analysis. Such errors in sentiment scores are common across libraries and packages for textual and sentiment analytics, and the general expectation is that the overall sentiment scores remain fairly accurate subject to availability of sufficient data. Considering this issue, we used two more lexicon-based python libraries (e.g., *Textblob* and *VADER*) to validate the sentiment score estimated using the *Syuzhet* package. However, future studies should consider fusing information from Twitter messages and sentiment propagation patterns to achieve better performance in analyzing Twitter data [64].
- Fourth, this study explored the public sentiment towards the COVID-19 vaccine in the US and investigated the spatial and temporal changes in the sentiments. However, a study evaluating the key reasons for positive and negative sentiments of the people is necessary for the targeted intervention of the authority to increase vaccine uptake.

VII. POLICY IMPLICATIONS

Based on the sentiment analysis and our review of multiple surveys, we posit that local, state and federal governments can leverage our PSS framework driven analysis to influence policy initiatives for improving COVID-19 vaccination rates. Our analysis studied sentiment changes at the national and state levels and corresponding potential policy implications are stated as below.

- First, there is a significant opportunity for governments to improve vaccine distribution and delivery and perceptions surrounding the same. Our Twitter sentiment analysis and our review of associated information and multiple related surveys, demonstrates that the vaccine distribution and delivery process is one of the prominent reasons why people are unhappy with regards to vaccination. Therefore, federal and state governments can adapt their policies to both improve the distribution system and process, and also improve communications about the distribution and delivery process to address public concerns.
- Second, from the temporal sentiment analytics studying public sentiment changes from early Feb'21 to late March'21, we have observed that simply giving more time did not help to improve the public perceptions about COVID-19 vaccination. When vaccines became available to an increasing majority of people in the US, it did not proportionately improve the vaccination rates [65]. Hence, the findings of our study imply that it would be beneficial for federal and state governments to take proactive actions to specifically attract vaccine hesitant sections of the population to get vaccinated. For example, the state of West Virginia offered \$100 savings bonds to 16-35 years vaccinated adults. The state of Maryland offered \$100 to all the vaccinated state employees. In Connecticut, New Jersey and Washington D.C., there are free-drink offers for the vaccine takers [66]. An increase in such initiatives has the potential to increase the COVID-19 vaccination rates.
- Third, from our application of the PSS framework and our thematic review of tweets, it is evident that people have hesitation and fear about the side effects of vaccination, and have doubts about the effectiveness of the vaccines, leading to increasing negative sentiment from early Feb to late March'21. This scenario creates a need for governments to launch nationwide and statewide public awareness programs to educate people about vaccines, their side-effects and effectiveness in a transparent and professional manner. It may be possible to learn from states where public sentiments show higher proportions of joy, trust and anticipation towards vaccines. Exceptional success stories need to be identified and disseminated through mass media to strengthen the positive sentiment towards vaccination. Moreover, sentiment analysis from this research can be associated with socioeconomic factors for further research to know how culture, income level, education level, and religious belief affect sentiment and public opinion variance toward COVID-19 vaccines.
- Fourth, some studies (e.g., [38]) have observed a higher vaccine hesitancy among specific sections of the society (e.g., Black/Hispanic population, low educated individuals, rural residents). This appears to be supported by linguistic patterns in certain tweets with negative sentiment. Thus, targeted and multi-faceted efforts could be effective in increasing the acceptance of COVID-19 vaccines by these sections of society.
- Fifth, An important insight from our application of sentiment analytics and textual thematic reviews, is that it has become evident that some sections of people are hesitant to take COVID-19 vaccines because of lack of trust based on the knowledge of the rapid development and approval of the COVID-19 vaccine, whereas other vaccines like Measles, Mumps etc; historically took significantly longer times to be developed, tested and approved. Governments and vaccine manufacturing pharmaceutical companies need to proactively develop policies to publicise the reasons and the science behind the relatively rapid development and emergency approval of the use of COVID vaccines via social media and supporting forms of mass media communications. This could increase positive sentiment and public trust towards the vaccine.

- Sixth, governmental and corporate policies for news media and social media driven publicizing of vaccinations administered to senior government officials, CEOs of companies, and such other public and influential figures in society (e.g., celebrities) could positively influence and motivate people to accept the COVID-19 vaccine [35]. We observed the presence of positive opinions on this topic in our textual thematic analysis of tweets.
- Seventh, our analysis highlights the importance of using the PSS framework for generating insights on the collective public sentiment towards COVID-19 vaccines at both the federal as well as at the state levels. We see varying changes between the national and local level aggregate public sentiment as shown in Figs. 7, 9 and 11. Furthermore, the thematic analysis indicates that some sections of the population which are vaccine hesitant have a high degree of concern for their health, as opposed to other sections who are vaccine hesitant and display little concern regarding COVID-19 health issues. We posit that governments must therefore accommodate state and local level needs of the people, and also proactively cater to the opportunity to develop policies that support all sections of the citizenry, even if vaccine hesitant, in helping them safeguard their health by boosting natural immunity and by using methods they are most willing to accept.

VIII. CONCLUSION AND FUTURE WORK

In this work, we have analyzed public sentiment in the US about COVID-19 vaccine using Twitter data. The findings from Twitter data analytics were verified and validated by the actual vaccination data from US CDC and recently performed Nationwide Household Pulse Survey (HPS) conducted by the US Census Bureau. We have analyzed the progression of public sentiment from early February to late March of 2021. We have found positive sentiment to be dominant over negative sentiment about COVID-19 vaccine in both Feb'21 and March'21. However, interestingly, in the sentiment progression from Feb to March, we noticed that positive sentiment declines and neutral sentiment increases, along with an increase in negative sentiment. We have demonstrated these findings both with descriptive analytics and statistical analysis. We have used the PSS (Public Sentiment Scenario) framework to structure our hypotheses for statistical analysis. Finally, the hypotheses are tested using Proportion test, Exact Binomial test and guided by the Wilcoxon signed rank test, demonstrating that our findings are non-trivial.

In our state-wise sentiment analysis, we have found that sentiment score differs from one state to another. We have also noticed that positive state-wise sentiment score has a strong correlation with the actual vaccination in that state. Based on our sentiment analysis, and extensive review of vaccination rates and survey data, we have observed that simply giving more time may not attract the "low-hanging fruit" type Americans to get vaccinated. Rather, the agencies and organizations of local, state and federal governments

need to make proactive efforts and provide incentives to attract vaccine hesitant Americans to accept the vaccine. Finally, we have provided valuable inputs for future public policies to increase the vaccine uptake and help reach the nation's much cherished herd immunity goal.

The current study was conducted in a time sensitive environment with Twitter data to gauge public perceptions of the COVID-19 vaccine. However, future analysis of vaccine sentiment based on information collected from multiple sources (e.g., Facebook, LinkedIn, Gab, various public online discussion boards, etc.) could cover a larger audience from different and multiple segments [34], [67]. Future research can leverage the current study along with information from multiple social media platforms in diverse formats (e.g., text, blogs, images, comments etc.) to develop a wide-ranging array of scenarios for sentiment analytics, and thus expand the application of the PSS framework.

Considering the enormous computational ability of the machine learning based techniques to handle complex and multifaceted problems, future studies should apply machine learning based algorithms to classify tweets accurately that can symbolize a real-world context [17], [68], [69]. This study employs the PSS framework, and provides a strong basis for formalizing public sentiment driven influence on policy formation and implementation. This is very valuable, and can lead to applied solutions, especially in conjunction with automated artificially intelligent information generation processes which, for example, can be used to machine-generate intelligent tweet responses and build positive consensus, clarify misinformation and boost supportive sentiment for the vaccine [70]–[72].

REFERENCES

- [1] CDC, "US COVID-19 cases and deaths," CDC, USA, 2021, accessed on: 2021-05-15. [Online]. Available: <https://covid.cdc.gov/covid-data-tracker/>
- [2] Worldometers, "COVID-19 Coronavirus Pandemic." Worldometer, USA, 2021, accessed on 2021-05-14. [Online]. Available: <https://www.worldometers.info/coronavirus/#countries>
- [3] S. Bell, R. Clarke, S. Mounier-Jack, J. L. Walker, and P. Paterson, "Parents' and guardians' views on the acceptability of a future COVID-19 vaccine: A multi-methods study in England," *Vaccine*, vol. 38, no. 49, pp. 7789–7798, 2020.
- [4] J. Samuel, G. Ali, M. Rahman, E. Esawi, and Y. Samuel, "COVID-19 public sentiment insights and machine learning for tweets classification," *Information*, vol. 11, no. 6, 2020, Art. no. 314.
- [5] J. Samuel, M. M. Rahman, G. G. M. N. Ali, Y. Samuel, A. Pelaez, P. H. J. Chong, and M. Yakubov, "Feeling Positive About Reopening? New Normal Scenarios From COVID-19 US Reopen Sentiment Analytics," *IEEE Access*, vol. 8, pp. 142 173–142 190, 2020.
- [6] "Herd immunity," Wikipedia, accessed on: 2021-05-14. [Online]. Available: https://en.wikipedia.org/wiki/Herd_immunity
- [7] F. P. Polack, S. J. Thomas, N. Kitchen, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. Pérez Marc, E. D. Moreira, C. Zerbini, R. Bailey, K. A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R. W. Frenck, L. L. Hammit, Türeci, H. Nell, A. Schaefer, S. Ünal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Şahin, K. U. Jansen, and W. C. Gruber, "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine," *New England Journal of Medicine*, vol. 383, no. 27, pp. 2603–2615, 2020. [Online]. Available: <https://doi.org/10.1056/NEJMoa2034577>
- [8] L. R. Baden, H. M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S. A. Spector, N. Rouphael, C. B. Creech, J. McGettigan, S. Khetan, N. Segall, J. Solis, A. Brosz, C. Fierro, H. Schwartz, K. Neuzil,

- L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B. S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. Deng, H. Zhou, S. Han, M. Ivarsson, J. Miller, and T. Zaks, "Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine," *New England Journal of Medicine*, vol. 384, no. 5, pp. 403–416, 2021. [Online]. Available: <https://doi.org/10.1056/NEJMoa2035389>
- [9] A. Apple, T. Azimi, and J. Cordina, "COVID-19 vaccine: Are US consumers ready?" McKinsey & Company, USA, December 2020, accessed on 2021-05-14. [Online]. Available: <https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/covid-19-vaccine-are-consumers-ready>
- [10] R. Robbins, "Millions Are Skipping Their Second Doses of Covid Vaccines," *The New York Times*, USA, April 2021, accessed on 2021-05-14. [Online]. Available: <https://www.nytimes.com/2021/04/25/business/covid-vaccines-second-doses.html>
- [11] K. Goulet, A. Jain, L. Koslow, and M. Gjaja, "The COVID-19 US Vaccine Sentiment Series: A Looming Slowdown in Demand," BCG, USA, April 2021, accessed on 2021-05-14. [Online]. Available: <https://www.bcg.com/en-us/publications/2021/covid-19-us-vaccine-sentiment-series>
- [12] CDC, "COVID-19 vaccinations in the United States," CDC, USA, 2021, accessed on: 2021-05-14. [Online]. Available: <https://covid.cdc.gov/covid-data-tracker/vaccinations>
- [13] U.S. Census Bureau, "Week 28 Household Pulse Survey: April 14 – April 26," 2021, accessed on 2021-05-14. [Online]. Available: <https://bit.ly/3vToIFK>
- [14] A. Borriello, D. Master, A. Pellegrini, and J. M. Rose, "Preferences for a COVID-19 vaccine in Australia," *Vaccine*, vol. 39, no. 3, pp. 473–479, 2021.
- [15] M. Rahman, J.-C. Thill, K. C. Paul *et al.*, "COVID-19 Pandemic Severity, Lockdown Regimes, and People's Mobility: Early Evidence from 88 Countries," *Sustainability*, vol. 12, no. 21, 2020, Art. no. 9101.
- [16] L. Coudeville, G. Gomez, O. Jollivet, R. Harris, E. Thommes, S. Druelles, A. Chit, S. Chaves, and C. Mahé, "Exploring uncertainty and risk in the accelerated response to a COVID-19 vaccine: Perspective from the pharmaceutical industry," *Vaccine*, vol. 38, no. 48, p. 7588–7595, 2020.
- [17] M. Rahman, K. C. Paul, M. Hossain, G. NawazAli, J.-C. Thill *et al.*, "Machine Learning on the COVID-19 Pandemic, Human Mobility and Air Quality: A Review," *IEEE Access*, 2021, Early Access version on webpage at <https://ieeexplore.ieee.org/abstract/document/9427494>.
- [18] M. M. Rahman, G. M. N. Ali, X. J. Li, J. Samuel, K. C. Paul, P. H. Chong, and M. Yakubov, "Socioeconomic factors analysis for COVID-19 US reopening sentiment with Twitter and census data," *Heliyon*, 2021, Art. no. e06200.
- [19] L. C. Karlsson, A. Soveri, S. Lewandowsky, L. Karlsson, H. Karlsson, S. Nolvi, M. Karukivi, M. Lindfelt, and J. Antfolk, "Fearing the disease or the vaccine: The case of COVID-19," *Personality and individual differences*, vol. 172, 2021, Art. no. 110590.
- [20] A. R. Mercadante and A. V. Law, "Will They, or Won't They? Examining Patients' Vaccine Intention for Flu and COVID-19 using the Health Belief Model," *Research in Social and Administrative Pharmacy*, 2020.
- [21] M. Schwarzingler, V. Watson, P. Arwidson, F. Alla, and S. Luchini, "COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics," *The Lancet Public Health*, vol. 6, no. 4, pp. e210–e221, 2021.
- [22] R. E. Glover, R. Urquhart, J. Lukawska, and K. G. Blumenthal, "Vaccinating against covid-19 in people who report allergies," 2021.
- [23] G. Troiano and A. Nardi, "Vaccine hesitancy in the era of COVID-19," *Public Health*, 2021.
- [24] N. E. MacDonald *et al.*, "Vaccine hesitancy: Definition, scope and determinants," *Vaccine*, vol. 33, no. 34, pp. 4161–4164, 2015.
- [25] D. A. Salmon, M. Z. Dudley, J. M. Glanz, and S. B. Omer, "Vaccine hesitancy: causes, consequences, and a call to action," *Vaccine*, vol. 33, pp. D66–D71, 2015.
- [26] C. S. Wiysonge, D. Ndwandwe, J. Ryan, A. Jaca, O. Batouré, B.-P. M. Anya, and S. Cooper, "Vaccine hesitancy in the era of COVID-19: could lessons from the past help in divining the future?" *Human vaccines & immunotherapeutics*, pp. 1–3, 2021.
- [27] M. Sallam, "Covid-19 vaccine hesitancy worldwide: A concise systematic review of vaccine acceptance rates," *Vaccines*, vol. 9, no. 2, 2021, Art. no. 160.
- [28] M. Thelwall, K. Kousha, and S. Thelwall, "Covid-19 vaccine hesitancy on English-language Twitter," 2021. [Online]. Available: <http://hdl.handle.net/2436/624008>
- [29] H. Lyu, J. Wang, W. Wu, V. Duong, X. Zhang, T. D. Dye, and J. Luo, "Social Media Study of Public Opinions on Potential COVID-19 Vaccines: Informing Dissent, Disparities, and Dissemination," *arXiv*, 2020, preprint on webpage at <https://arxiv.org/abs/2012.02165>.
- [30] H. Piedrahita-Valdés, D. Piedrahita-Castillo, J. Bermejo-Higuera, P. Guillem-Saiz, J. R. Bermejo-Higuera, J. Guillem-Saiz, J. A. Sicilia-Montalvo, and F. Machío-Regidor, "Vaccine Hesitancy on Social Media: Sentiment Analysis from June 2011 to April 2019," *Vaccines*, vol. 9, no. 1, 2021. [Online]. Available: <https://www.mdpi.com/2076-393X/9/1/28>
- [31] S. Thukraal, "COVID-19 vaccines: Social media user sentiments and strategies to encourage administration," *Clarivate*, April 2021. [Online]. Available: <https://clarivate.com/cortellis/article/covid-19-vaccines-social-media-user-sentiments-and-strategies-to-encourage-administration/>
- [32] V. Raghupathi, J. Ren, and W. Raghupathi, "Studying Public Perception about Vaccination: A Sentiment Analysis of Tweets," *International Journal of Environmental Research and Public Health*, vol. 17, no. 10, 2020. [Online]. Available: <https://www.mdpi.com/1660-4601/17/10/3464>
- [33] M. Salathé and S. Khandelwal, "Assessing Vaccination Sentiments with Online Social Media: Implications for Infectious Disease Dynamics and Control," *PLOS Computational Biology*, vol. 7, no. 10, pp. 1–7, 10 2011. [Online]. Available: <https://doi.org/10.1371/journal.pcbi.1002199>
- [34] A. Hussain, A. Tahir, Z. Hussain, Z. Sheikh, M. Gogate, K. Dashtipour, A. Ali, and A. Sheikh, "Artificial Intelligence-Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States: Observational Study," *Journal of medical Internet research*, vol. 23, no. 4, 2021, Art. no. e26627.
- [35] J. Griffith, H. Marani, and H. Monkman, "COVID-19 Vaccine Hesitancy in Canada: Content Analysis of Tweets Using the Theoretical Domains Framework," *Journal of medical Internet research*, vol. 23, no. 4, 2021, Art. no. e26874.
- [36] F. Eibensteiner, V. Ritschl, F. A. Nawaz, S. S. Fazel, C. Tsagkaris, S. T. Kulnik, R. Crutzen, E. Klager, S. Völkl-Kernstock, E. Schaden *et al.*, "People's Willingness to Vaccinate Against COVID-19 Despite Their Safety Concerns: Twitter Poll Analysis," *Journal of Medical Internet Research*, vol. 23, no. 4, 2021, Art. no. e28973.
- [37] J. B. Ruiz and R. A. Bell, "Predictors of intention to vaccinate against covid-19: Results of a nationwide survey," *Vaccine*, vol. 39, no. 7, pp. 1080–1086, 2021.
- [38] K. A. Fisher, S. J. Bloomstone, J. Walder, S. Crawford, H. Fouayzi, and K. M. Mazor, "Attitudes toward a potential SARS-CoV-2 vaccine: a survey of US adults," *Annals of internal medicine*, vol. 173, no. 12, pp. 964–973, 2020.
- [39] J. K. Ward, C. Alleaume, P. Peretti-Watel, V. Seror, S. Cortaredona, O. Launay, J. Raude, P. Verger, F. Beck, S. Legleye *et al.*, "The french public's attitudes to a future covid-19 vaccine: The politicization of a public health issue," *Social science & medicine*, vol. 265, p. 113414, 2020.
- [40] M. Detoc, S. Bruel, P. Frappe, B. Tardy, E. Botelho-Nevers, and A. Gagneux-Brunon, "Intention to participate in a covid-19 vaccine clinical trial and to get vaccinated against covid-19 in france during the pandemic," *Vaccine*, vol. 38, no. 45, pp. 7002–7006, 2020.
- [41] K. Wang, E. L. Y. Wong, K. F. Ho, A. W. L. Cheung, E. Y. Y. Chan, E. K. Yeoh, and S. Y. S. Wong, "Intention of nurses to accept coronavirus disease 2019 vaccination and change of intention to accept seasonal influenza vaccination during the coronavirus disease 2019 pandemic: A cross-sectional survey," *Vaccine*, vol. 38, no. 45, pp. 7049–7056, 2020.
- [42] K. Wang, E. L.-Y. Wong, K.-F. Ho, A. W.-L. Cheung, P. S.-Y. Yau, D. Dong, S. Y.-S. Wong, and E.-K. Yeoh, "Change of willingness to accept covid-19 vaccine and reasons of vaccine hesitancy of working people at different waves of local epidemic in hong kong, china: Repeated cross-sectional surveys," *Vaccines*, vol. 9, no. 1, p. 62, 2021.
- [43] A. L. Wagner, Z. Huang, J. Ren, M. Laffoon, M. Ji, L. C. Pinckney, X. Sun, L. A. Prosser, M. L. Boulton, and B. J. Zikmund-Fisher, "Vaccine hesitancy and concerns about vaccine safety and effectiveness in Shanghai, China," *American journal of preventive medicine*, vol. 60, no. 1, pp. S77–S86, 2021.
- [44] S. Kreps, S. Prasad, J. S. Brownstein, Y. Hswen, B. T. Garibaldi, B. Zhang, and D. L. Kriner, "Factors associated with US adults' likelihood of accepting COVID-19 vaccination," *JAMA network open*, vol. 3, no. 10, pp. e2025594–e2025594, 2020.
- [45] R. D. Goldman, S. R. Marni, M. Seiler, J. C. Brown, E. J. Klein, C. P. Cotanda, R. Gelernter, T. D. Yan, J. Hoeffe, A. L. Davis *et al.*, "Caregivers'

- willingness to accept expedited vaccine research during the COVID-19 pandemic: a cross-sectional survey," *Clinical Therapeutics*, vol. 42, no. 11, pp. 2124–2133, 2020.
- [46] A. Leng, E. Maitland, S. Wang, S. Nicholas, R. Liu, and J. Wang, "Individual preferences for COVID-19 vaccination in China," *vaccine*, vol. 39, no. 2, pp. 247–254, 2021.
- [47] P. L. Reiter, M. L. Pennell, and M. L. Katz, "Acceptability of a covid-19 vaccine among adults in the united states: How many people would get vaccinated?" *Vaccine*, vol. 38, no. 42, pp. 6500–6507, 2020.
- [48] A. C. Guelzo, "'public sentiment is everything': Abraham lincoln and the power of public opinion," 2014.
- [49] R. Y. Shapiro, "Public opinion and american democracy," *Public Opinion Quarterly*, vol. 75, no. 5, pp. 982–1017, 2011.
- [50] J. Höchtl, P. Parycek, and R. Schöllhammer, "Big data in the policy cycle: Policy decision making in the digital era," *Journal of Organizational Computing and Electronic Commerce*, vol. 26, no. 1-2, pp. 147–169, 2016.
- [51] J. Samuel, R. Holowczak, and A. Pelaez, "The effects of technology driven information categories on performance in electronic trading markets," *Journal of Information Technology Management*, vol. 28, no. 1-2, pp. 1–14, 2017.
- [52] R. C. Black, R. J. Owens, J. Wedeking, and P. C. Wohlfarth, "The influence of public sentiment on supreme court opinion clarity," *Law & Society Review*, vol. 50, no. 3, pp. 703–732, 2016.
- [53] J. Samuel, R. Holowczak, R. Benbunan-Fich, and I. Levine, "Automating discovery of dominance in synchronous computer-mediated communication," in *2014 47th Hawaii International Conference on System Sciences*. IEEE, 2014, pp. 1804–1812.
- [54] S. Li, Z. Liu, and Y. Li, "Temporal and spatial evolution of online public sentiment on emergencies," *Information Processing & Management*, vol. 57, no. 2, 2020, Art. no. 102177.
- [55] J. Samuel, "Information token driven machine learning for electronic markets: Performance effects in behavioral financial big data analytics," *Journal of Information Systems and Technology Management*, vol. 14, no. 3, pp. 371–383, 2017.
- [56] —, "Informatics in information richness: A market mover? an examination of information richness in electronic markets," *Samuel, J. & Pelaez, A., (2017). Informatics in Information Richness: A Market Mover*, 2017.
- [57] T. Rinker, "Calculate Text Polarity Sentiment," March 2019, accessed on 2021-05-14. [Online]. Available: <https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf>
- [58] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, May 2014. [Online]. Available: <https://ojs.aaai.org/index.php/ICWSM/article/view/14550>
- [59] O. Oyeboade and R. Orji, "Social Media and Sentiment Analysis: The Nigeria Presidential Election 2019," in *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2019, pp. 0140–0146.
- [60] The New York Times, "See how vaccinations are going in your county and state," 2021, accessed: 2020-05-7. [Online]. Available: <https://www.nytimes.com/interactive/2020/us/covid-19-vaccine-doses.html>
- [61] S. P. Jones, "Imperial college london big data analytical unit and yougov plc. 2020," *Imperial College London YouGov Covid Data Hub*, v1. 0, YouGov Plc, 2020. [Online]. Available: <https://bit.ly/3tvDqkt>
- [62] R. Nagar, Q. Yuan, C. C. Freifeld, M. Santillana, A. Nojima, R. Chunara, and J. S. Brownstein, "A case study of the New York City 2012-2013 influenza season with daily geocoded Twitter data from temporal and spatiotemporal perspectives," *Journal of medical Internet research*, vol. 16, no. 10, 2014, Art. no. e236.
- [63] Y. M. Çetinkaya, İ. H. Toroslu, and H. Davulcu, "Developing a twitter bot that can join a discussion using state-of-the-art architectures," *Social Network Analysis and Mining*, vol. 10, no. 1, pp. 1–21, 2020.
- [64] L. Wang, J. Niu, and S. Yu, "Sentidiff: Combining textual information and sentiment diffusion patterns for twitter sentiment analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 10, pp. 2026–2039, 2019.
- [65] S. Guttentag, "What's Causing Low COVID-19 Vaccine Uptake — Hesitancy, Lack of Access, or Both?," *GoodRx*, April 2021, accessed on: 2021-05-14. [Online]. Available: <https://www.goodrx.com/blog/low-covid-19-vaccine-uptake-causes/>
- [66] K. Hubbard, "Want Free Beer or a Chance at \$1 Million? Get Your COVID-19 Vaccine," May 2021, accessed on: 2021-05-15. [Online]. Available: <https://www.usnews.com/news/best-states/articles/2021-05-07/states-cities-and-companies-offer-incentives-to-get-covid-19-vaccine>
- [67] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," *Knowledge and Information Systems*, vol. 60, no. 2, pp. 617–663, 2019.
- [68] D. Stojanovski, G. Strezoski, G. Madjarov, and I. Dimitrovski, "Twitter sentiment analysis using deep convolutional neural network," in *International Conference on Hybrid Artificial Intelligence Systems*. Springer, 2015, pp. 726–737.
- [69] A. Hassan, A. Abbasi, and D. Zeng, "Twitter sentiment analysis: A bootstrap ensemble framework," in *2013 international conference on social computing*. IEEE, 2013, pp. 357–364.
- [70] M. D. Garvey, J. Samuel, and A. Pelaez, "Would you please like my tweet?! an artificially intelligent, generative probabilistic, and econometric based system design for popularity-driven tweet content generation," *Decision Support Systems*, vol. 144, p. 113497, 2021.
- [71] J. Samuel, M. Garvey, and R. Kashyap, "That message went viral?! exploratory analytics and sentiment analysis into the propagation of tweets," *arXiv preprint arXiv:2004.09718*, 2020.
- [72] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, "Bertscore: Evaluating text generation with bert," *arXiv preprint arXiv:1904.09675*, 2019.



G. G. MD. NAWAZ ALI (Member, IEEE) received his B.Sc. degree in Computer Science and Engineering from the Khulna University of Engineering & Technology, Khulna, Bangladesh in 2006, and the Ph.D. degree in Computer Science from the City University of Hong Kong, Hong Kong in 2013 with the Outstanding Academic Performance Award. He is currently working as an Assistant Professor with the Department of Applied Computer Science of the University of Charleston, West Virginia, USA. Prior to joining UCWV, he was a Postdoctoral Fellow with the Department of Automotive Engineering, The Clemson University International Center for Automotive Research (CU-ICAR), Greenville, South Carolina, USA from March 2018 to July 2019. From October 2015 to March 2018, he was a postdoctoral research fellow with the School of Electrical and Electronic Engineering of Nanyang Technological University (NTU), Singapore. He is a reviewer of a number of international journals including the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS AND MAGAZINE, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE INTERNET OF THINGS (IOT) JOURNAL, IEEE ACCESS, AND Wireless Networks etc. His current research interests include Vehicular Cyber Physical System (VCPS), wireless broadcasting, mobile computing, and network coding.



MD. MOKHLESUR RAHMAN has completed his Bachelor of Urban and Regional Planning (B.URP) from the Department of Urban and Regional Planning (URP), Bangladesh University of Engineering & Technology (BUET), Dhaka, Bangladesh in 2009. To compete in the professional arena and acquire advanced knowledge, he has completed an M.Sc in Urban Planning from the Department of Urban Planning and Design (DUPAD), University of Hong Kong (HKU), Hong Kong. After graduation from HKU, he started working at Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh as a Lecturer in December 2014 and promoted to Assistant Professor in February 2017. Currently, he is a graduate student and research assistant at the University of North Carolina at Charlotte, North Carolina, USA. His research interests include Transportation Planning and Engineering, Travel Behavior and Demand Modeling, Land Use Modeling, Climate Change and Disaster Management, Applications of GIS in Transportation and Environment, Econometric Modeling, and Data Analytics.



MD. AMJAD HOSSAIN (Member, IEEE) received B.Sc. degree in computer science and Engineering from Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh, in 2008. He completed his Ph.D. in Computer Science from Kent State University, Ohio, USA, in 2020. Hossain started working as a Lecturer in KUET immediately after his B.Sc. Later, he received a promotion to Assistant Professor and continued working in KUET until 2012. Currently, he is working as an Assistant Professor of computer and information sciences at Shepherd University, West Virginia, USA. His current research interest includes distributed systems, multimedia computing and networking, ML, and Image Processing.



MD. SHAHINOOR RAHMAN is an assistant professor at the Department of Earth and Environmental Sciences of New Jersey City University (NJCU), New Jersey, USA. He completed his Ph.D. in Earth Systems and Geoinformation Sciences from George Mason University (GMU), Virginia, USA in December 2019. He is the recipient of the Outstanding Ph.D. Student Award 2020 from the Department of Geography and Geoinformation Science at Mason for his excellence in the studies, prolific track record of scholarly publications. A former Assistant Professor of the apex engineering university in Dhaka, Bangladesh, Bangladesh University of Engineering and Technology (BUET), he completed his undergraduate in Urban and Regional Planning from BUET and masters in Regional Development Planning and Management jointly offered by TU-Dortmund, Dortmund, Germany and Universidad Austral de Chile, Los Ríos, Chile. He primarily specializes in the application of remote sensing (RS) and geographic information science (GIS) to examine the cause of—and innovative solutions to—some of our world's most pressing environmental challenges including climate change impact on soil and water, agricultural damage, coastal vulnerability, natural hazard risks, environmental degradation by urban expansion, and urban heat island effect. He utilizes data mining, ML, and spatial analytics approaches in his research. His broad goal is to utilize advanced geospatial technologies for the betterment of our society and environment. He is a reviewer of many reputed journals in remote sensing and GIS fields. He is currently serving as a special issue editor of MDPI Land.



KAMAL CHANDRA PAUL (Graduate Student Member, IEEE) is a PhD student at the department of Electrical and Computer Engineering, the University of North Carolina at Charlotte (UNCC), North Carolina, USA. He received his bachelor degree in Electrical and Electronic Engineering (EEE) from Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh in 2009. He obtained his MS degree in Electrical Engineering from the Ingram School of Engineering, Texas State University, San Marcos, Texas, USA in 2018. He served as a faculty at the department of EEE, International University of Business Agriculture and Technology (IUBAT), Dhaka, Bangladesh. He also worked as a lecturer in the department of EEE at the World University of Bangladesh, Dhaka, Bangladesh. His research interest includes AC and DC arc fault detection, data science, application of ML in power electronics, and renewable energy.



JEAN-CLAUDE THILL is a Knight Foundation Distinguished Professor of Public Policy in the Department of Geography and Earth Sciences at the University of North Carolina at Charlotte, North Carolina, USA. He is an Affiliate Faculty of the School of Data Science and holds graduate degrees from the Catholic University of Louvain, Louvain-la-Neuve, Belgium. Dr. Thill is an urban economic geographer whose research focuses on urban and regional transportation and mobility issues, processes of urbanization, regional science, and geospatial data science. His recent research focuses on modeling past urbanization trajectories of the Charlotte metropolitan area (land use transformation, local and regional drifts in neighborhood quality of life), livability, innovation, and mobility in Charlotte, other US cities, and Chinese cities, and accessibility to public services. He has more than 30 years of teaching and research experience in transportation planning and spatial modeling. His current and recent grants have been awarded from the National Science Foundation, Department of Defense, Department of Energy, North Carolina Department of Transportation, World Bank, and more. He is the 2012 recipient of the Edward L. Ullman Award for Significant Contributions to Transportation Geography, awarded by the American Association of Geographers. He has also received the Hirotada Kohno Award for Outstanding Service of the Regional Science Association International in 2012. Since 2013, he is a Fellow of the Regional Science Association International. He is honored to have been appointed High-End Foreign Expert in the School of Applied Economics, Renmin University of China (Ministry of Science and Technology), in Beijing, P.R. China.



JIM SAMUEL is Associate Professor and Director, Business Analytics & Applied Artificial Intelligence, at the School of Business, University of Charleston. He received his Ph.D. degree in Information Systems from Baruch College - Graduate Center, City University of New York, and has MBA (Finance) and M.Arch. degrees. Dr. Samuel's primary research covers human intelligence and artificial intelligence interaction, information philosophy, textual analytics and NLP, applied machine learning and socioeconomic implications of artificial intelligence. He has received multiple academic awards, is on editorial boards, and leads multiple analytics and applied artificial intelligence initiatives, involving corporate consulting and global research collaborations.

...